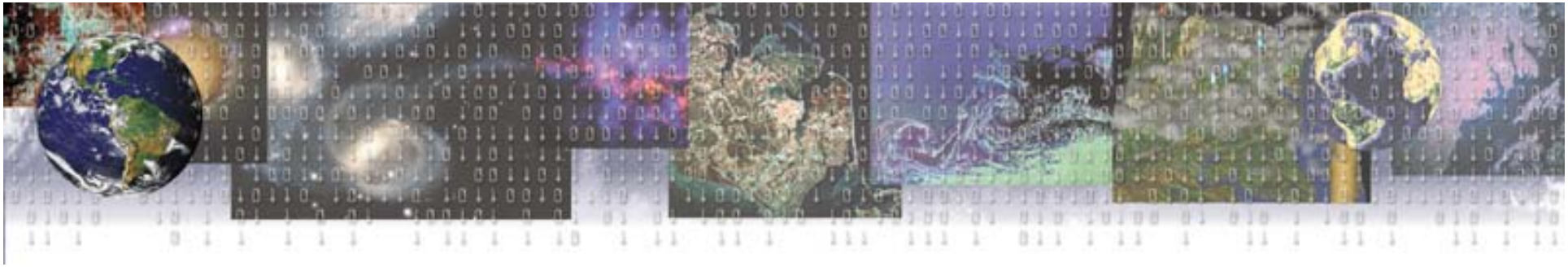
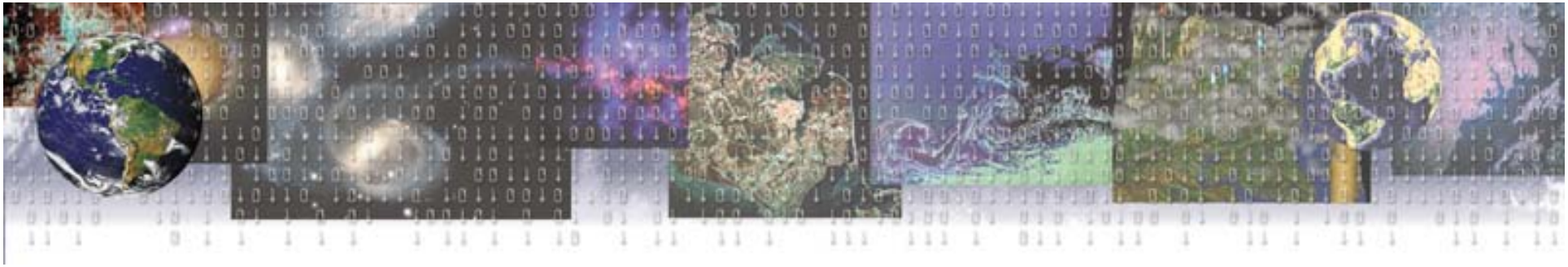


ECHO Development Activities

ECHO

Global Science & Technology, Inc.



Comparison of ECHO 4.5 to ECHO 5.0

ECHO

NASA'S MISSION TO PLANET EARTH
EARTH PROBES
DATA INFORMATION SYSTEM
EOS
EARTH OBSERVING SYSTEM

Global Science & Technology, Inc.

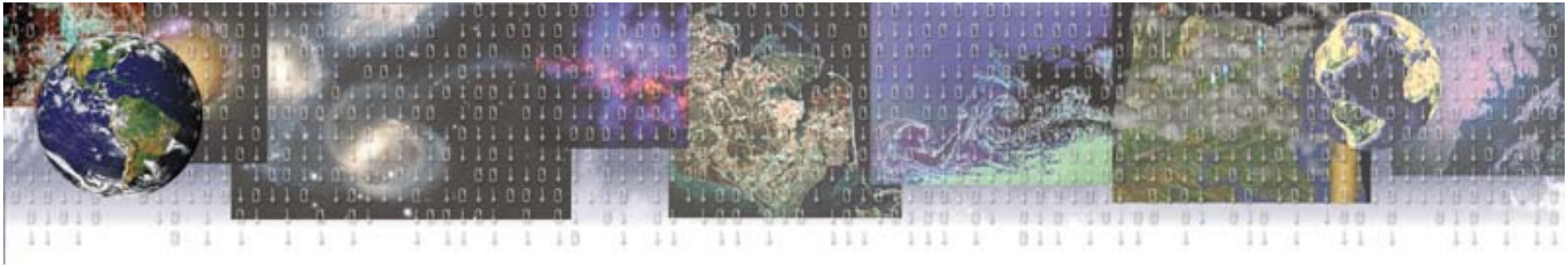
NASA

A Word About Iterative Development

- **Remember that ECHO is developed incrementally**
- **In each release:**
  - New functions may be added
  - Existing functions may be improved

Change in ECHO Unique ID

- **Known as ECHO Item ID**
- **Changed from numeric only field to a string field**



Roles in ECHO

-or-

How to have a split personality and maintain your API…

(A brief description of the roles capability added into ECHO Version 5)

Global Science & Technology, Inc.

4.5 System Baseline

- **Providers have separate accounts from registered users**
  - One account/password is shared by all provider personnel responsible for managing ECHO
- **Guests do not log in**

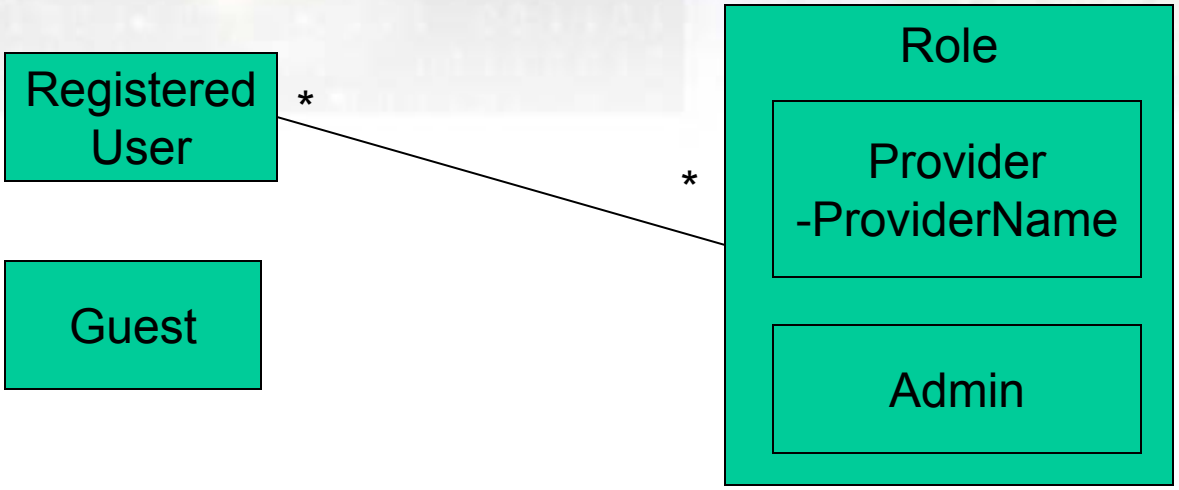
Two Types of Users

- **As of ECHO Version 5.0, there are two types of users**
  - Guest
    - This is a user of the system who does not call the login method of the Session Manager
    - This user is not guaranteed (or in most cases allowed) persistence of items in ECHO beyond their orders
  - Registered User
    - This is a user who has successfully called the login method of the Session Manager with their user id and password
    - This user can save and recall queries, result sets, persist account information, access a list of orders submitted or created but not submitted, etc.
- **Previously, there was a Provider User as well, but they may no longer log into the system**

Registered Users May Have Roles

- **Currently, there are two possible types of roles that may be given to a Registered User**
  - Note: A Registered User does not HAVE to have any roles
  - Administrator
    - A User with this role can access additional transactions that allow them to act as any user, delete users, and allow them to see any order. It does not allow them access to data that has been hidden via the DMS unless they act as a user that has access.
  - Provider
    - A User with this role can access additional transactions that allow them control over the Provider Account since no one can log in as the Provider anymore.
      - Provider Account Service
      - Provider Order Management Service
      - Data Management Service

Role



- A Registered User can have many Roles, including Administrator, and Provider for any number of Registered Providers
- A Guest User cannot have any Roles

Administrator Role

- **Once logged in, a User with this role acts as them self until they do a change context operation.**
- **If the user performs a change user context, then future transactions will operate as if the administrator is the user they changed context to**
  - The ECHO log files will show both the administrator's login name and the user they are acting as so that a record is maintained of who actually performed a function
- **If the administrator user performs a change provider context, then they will be allowed to access Provider only transactions, and will act as the specified Provider**
  - An admin user can change to any Provider

Provider Role

- **A User may have multiple Provider Roles, one for each Provider that they represent**
  - For instance, EDC has two separate data centers that they operate which may share personnel
- **Once a user has any Provider Role, then they are allowed to access additional service transactions**
  - If a user has exactly one Provider Role, then ECHO will use that Provider whenever a Provider Transaction is executed
  - If a user does not have exactly one Provider Role, then they must run the SetProviderContext method on the Session Manager to let ECHO know which Provider they are acting on the behalf of when they execute a Provider Transaction
    - Note: If a User has the Admin role, they can use SetProviderContext to become any Provider. If they have exactly one Provider Role in addition to the Admin Role, then they will not need to call SetProviderContext to act on behalf of that Provider. If they have more than one Provider Role, then they must call SetProviderContext

The Session Manager

- **The Session Manager is the object that any ECHO client talks to**
  - It maintains what user is currently using it
- **Session Manager has several methods:**
  - Login(UserID,Password)
    - Establishes which user is associated with the session
  - Logout()
    - Become a guest user again
  - Perform(XMLString):XMLString
    - How Service Transactions are called
  - Identify(String)
    - Associates a string that represents the client (e.g. EDG v3.8) that gets logged with the user information
  - SetProviderContext(ProviderName)
    - Allows a user that has multiple Provider Roles to select which Provider they are representing
  - SetUserContext(UserID)
    - Allows an administrator user to select which user they are acting as

Metadata Update Capability

- **Currently in 4.5, ECHO uses the ingest mechanism to insert, update and delete both collection and granule metadata**
  - Each collection or granule is updated as a whole
- **The Metadata Update capability allows certain fields to be updated independently of the rest of the granule**
  - In 5.0, this works for OnlineURL (for Data Pools) and QA flags only
  - The ECHO last update flag is updated appropriately
  - The Provider last update flag is updated according to the request

Catalog Service Changes

- **GetMetadata transaction**
  - Used to get metadata from ECHO without having to first do a query
  - Requires the ECHO unique identifier for a granule or collection
- **Query transaction**
  - New result type: ITEM_IDS
  - Returns only the ECHO unique identifier
  - No result set is created
- **Explain Search Parameter**
  - Also referred to as dynamic valids
  - Ability to find out what valid values exist in the database for a particular query attribute
    - Can be focused using a query

Order Entry Service

- **Set Authentication Key added**
  - Allows an authentication key to be added to a provider order
  - Not currently used when communicating with provider???
- **PresentOptionDefinitionsForProviderOrderRequest (ProviderOrderID) changed**
  - Changed to: PresentOrderOptionDefinitionsForProviderRequest (ProviderID)
- **SetOptionSelectionsForProviderOrderRequest (ProviderOrderID, OptionSelection+) changed**
  - Changed to: SetProviderOptionSelectionsForProviderOrderRequest (ProviderOrderID, OptionSelection+)
- **SetProviderOptionSelectionsForProviderOrderRequest (ProviderOrderID, OptionSelection+) added**
  - Added to allow checking what order options have already been set at the provider order level

OES Continued

- **Submit Order modified**
  - 4.5: Notification Mechanism: none or email
  - 5.0: Notification Level: Verbose, Detail, Info, Critical, None
    - Verbose: All state changes, status updates, and error messages
    - Detail: State changes and error messages
    - Info: State changes to closed or cancelled, and error messages
    - Critical: Error messages or rejection
    - None: No email sent
  - A User Preference has been created to be the default Notification Level (default for guest is Verbose)

Provider Account Service

- **Added role management functions**
  - Grant Provider Access
  - Revoke Provider Access
- **Present Provider Info**
  - Now returns the Provider ID needed to query the system

Provider Order Management Service

- **The infrastructure of ECHO was improved to allow the use of ORs in the API**
- **Several functions needed either a ProviderTrackingID or an OrderID**
  - They have been updated so that the OR is used appropriately
- **PresentOpenOrder, PresentOpenOrderSummary, PresentClosedOrder and PresentClosedOrderSummary changed to allow specification of order state as a filter**

Subscription Service

- **ListPausedSubscriptions**
  - Allows a user to list any of their subscriptions that they have previously paused
- **FTP defaults to passive mode**

User Account Service

- Changed CancelOrder to use the new API OR capability to express OR condition between Order ID and Provider Order ID
- ChangeUserPassword modified to allow an administrator to change a user's password without knowing the current password
- ListRoles function added to identify all roles granted to a user
- Added UserName to PresentUserInformation response

New User Preferences

- **Default notification level for orders**
- **Default shipping address**
  - Clients can use this to reference the address book for the address to use
- **Default billing address**
- **Default contact address**

Administration Service

- **Added administration capabilities to the system**

Provider Profile Service

- **ListAllProviders**
  - Allows a user to find out what Providers participate in the system
- **PresentProviderProfiles**
  - Gives access to Provider Contact information, description of holdings, spatial projection type and organization name
- **PresentProviderSupportedTransactions**
  - List of transactions that a provider actually participates in
  - Order, Quote, Cancel

New Provider Policies

- **Configuration of ECHO-> Provider communication parameters**
  - What server and port to talk to
  - What protocol to use (ODL or SOAP)
  - What transactions are supported (Quote, Order, Cancel)

The Facade

- **Set of Java classes that encapsulate all communication with ECHO**
- **Client developer only needs to use XML when creating queries, and interpreting results**
- **Utilizes SOAP interface to ECHO**
  - No knowledge of SOAP needed

New GUIs

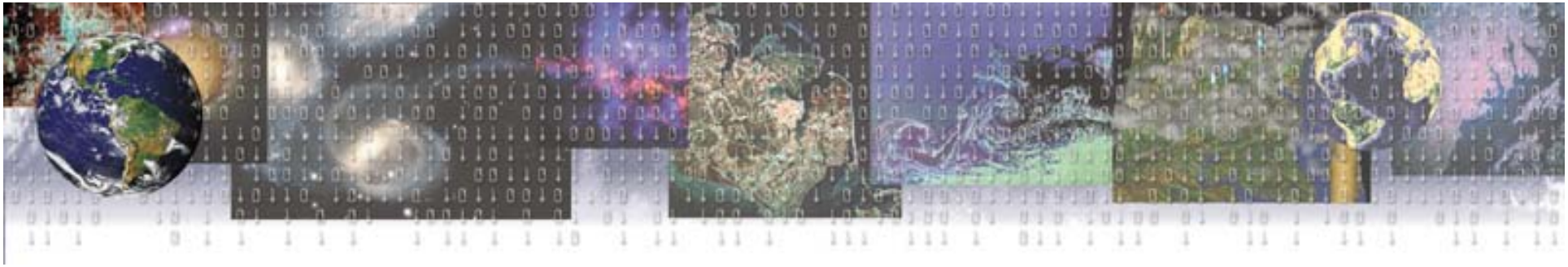
- **PUMP**
  - Access to provider account information, group management, ACL management, role management, and user account information
  - Creation of registered users, provider applications
- **Subscription Service GUI**
  - Management of metadata subscriptions
- **Provider Application GUI**
  - Convenient place to fill out provider application form
- **Provider Policies GUI**
  - Convenient place to discover and set provider policies

GUI Modifications

- **All GUIs use the façade**
- **Added user and provider context setting capabilities to test harness**
- **Added ability to point GUIs at any ECHO server**
  - This accommodates test installations of ECHO in addition to the operational one

ECHO 5.5 Overview

- **Services**
- **Inspection Functions**
- **Ingest Improvements**
- **Collection Based Search**
  - Backtrack Algorithm
- **DMR Changes**



Access Control Lists in ECHO

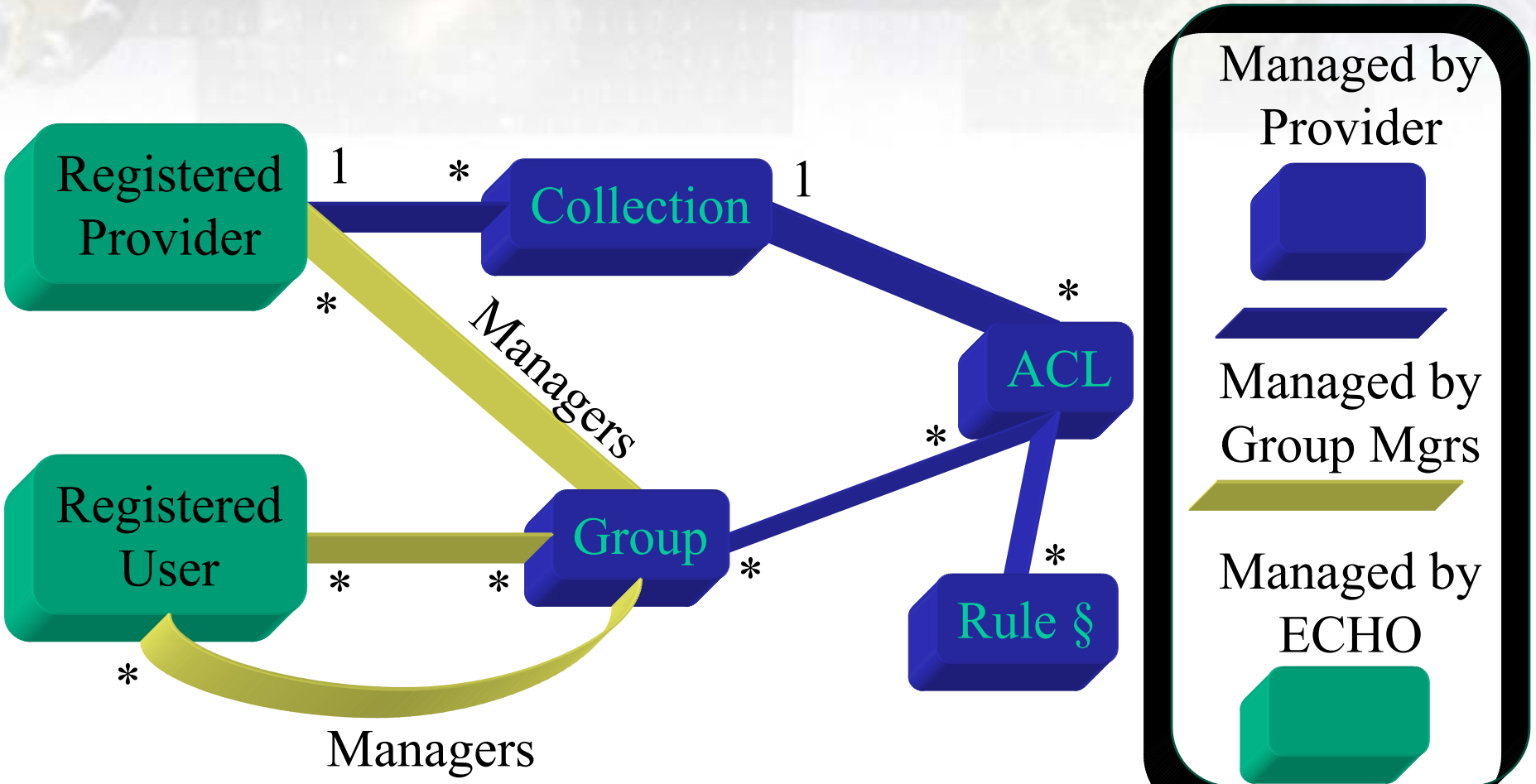
ECHO

Global Science & Technology, Inc.

Purpose of ACLs

- **Originally ECHO was intended to be a clearinghouse of public metadata**
  - ECHO can now also serve as a clearinghouse for restricted (private) metadata
- **Desire in provider community to use a single system to represent their metadata**
- **Also a desire to leverage user interfaces developed for ECHO for those who have access to restricted (private) metadata.**

Metadata Visibility – Conceptual Model



§Rule determines if collection is hidden and if not, which granules are hidden

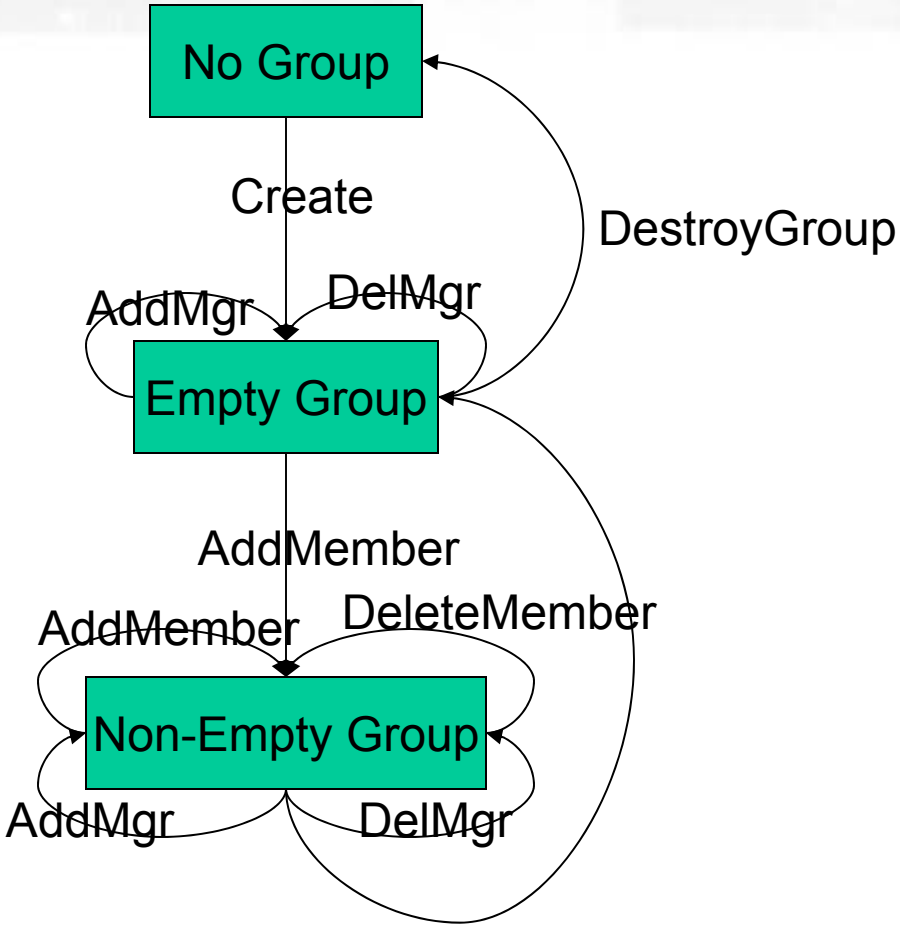
Key Concepts

- **Managers**
  - A list of users and/or providers
  - Can add or remove other managers
  - Can add or remove members of the group
  - Can contact the members or managers
- **Members**
  - A list of users and/or providers
  - Can contact other members
- **Real use: Metadata Visibility Permissions**

Why do we have groups?

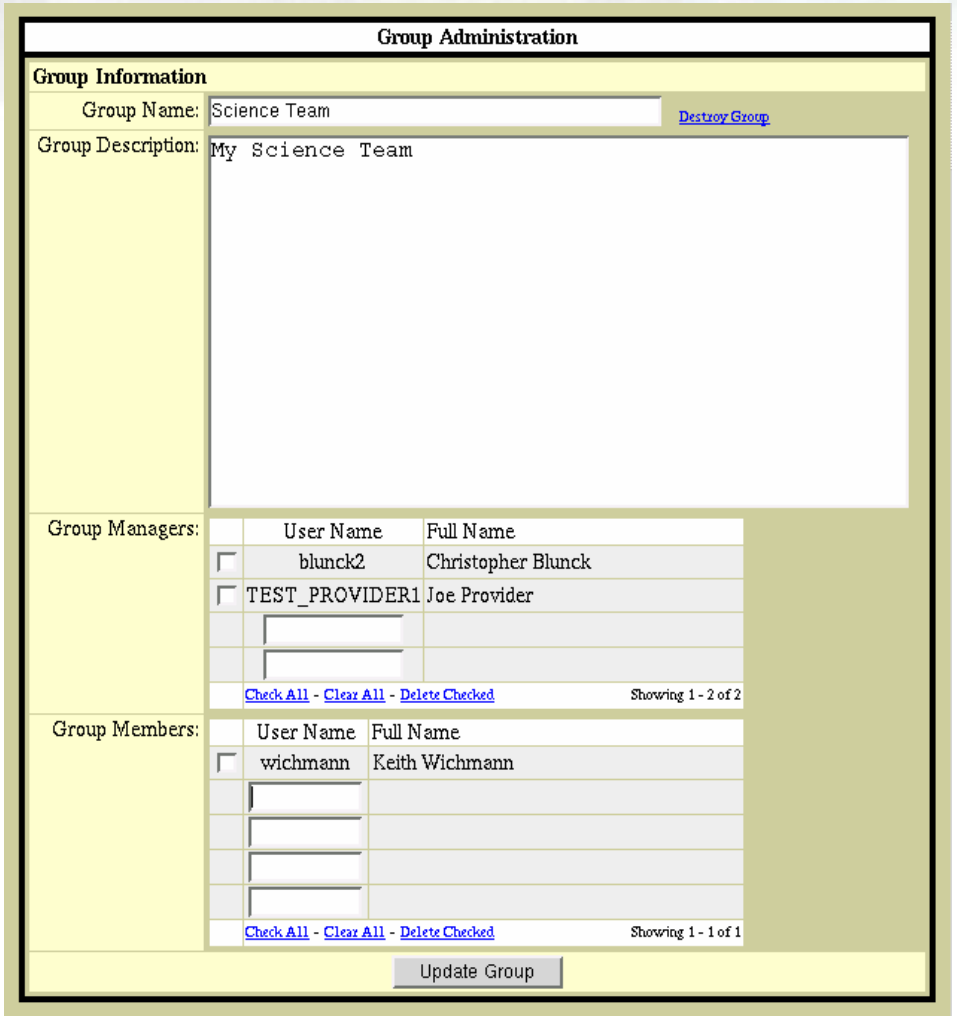
- **Mechanism for aggregating users**
- **Communication**
- **Metadata Visibility**

Group State Machine



Group Management Service GUI

- **Allows customers to manage groups**
  - Group Managers
  - Group Members
- **Managers can also contact group members**



Group Caveats

- **There can be multiple managers**
- **Any manager can add or remove other managers**
- **Managers are not automatically members of the group**
- **An owner has no control over the group, it is simply an attribute of the group that stores who originally created the group**
  - The owner will also be the first manager of a group if the group was created without specifying a manager

Data Management Service

- Used by providers to manage access control to their metadata and ordering processes
- Composed of Rules, Conditions and Groups
- Rules and Conditions combined are used to specify subsets of a collection to which to apply access control

ACL Approach

- **When do changes to ACLs take effect?**
  - In short, immediately (no batch updates needed)
  - The system is designed around the concept of provider control, and a provider's update is honored immediately in all ECHO transactions
- **Access Control Lists are honored at time of execution allowing a provider to make updates that will apply to existing result sets and orders**
  - This allows the provider to introduce new restrictions, remove existing restrictions, or arrange for a group to be updated
- **Certain types of restrictions are time-based and will expire automatically**
  - An ACL will allow access to a previously restricted result set when a temporal restriction expires
  - Does not require a provider to take an action

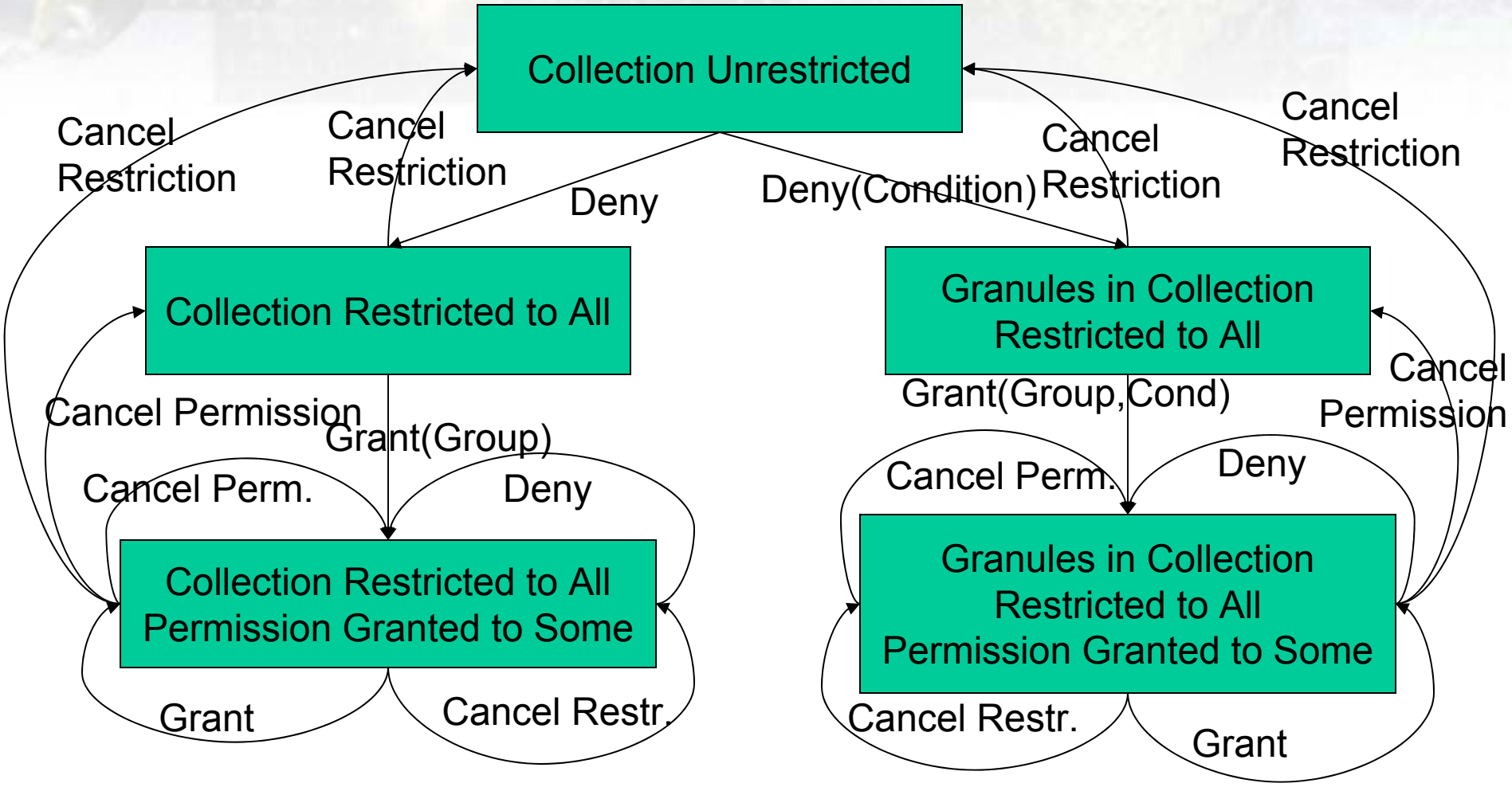
What access is being controlled?

- **Viewing**
  - Catalog Service and Subscription Service will prevent the presentation of controlled metadata
  - CAVEAT: Once a copy of the metadata is extracted from ECHO, it can be shared at will circumventing ECHO's access control functions exactly as is possible currently by ordering provider data and sharing with others
  - CAVEAT: Browse URLs that are part of the ECHO metadata are controlled in the sense that they are not presented if the metadata is not visible, but access is not checked for every client that accesses the URL directly
- **Ordering**
  - Order Entry Service prevents
    - Creation of orders with a restricted item
    - Adding restricted items to an order
    - Quoting orders with a restricted item
    - Submitting an order with a restricted item
- **Browse**
  - Browse has been discussed as being a separately controlled item, but is currently considered too complex

What is the granularity of control?

- **Collections**
  - Metadata Visibility: A collection's entire metadata description as well as all contained granules
    - CAVEAT: In the future, we plan to expose the name of the collection regardless of its access control state
  - Order: A collection or the granules within the collection
- **Granules**
  - An individual granule can be controlled by naming its Provider GranuleUR
  - This is discouraged except when it is very few granules affected
- **Subsets of Granules within a Collection**
  - Frequently it is useful to specify groups of granules within a collection by some common attribute to protect
  - Subsets can be performed based on Time of Acquisition
    - Fixed time periods or rolling time periods can be used
    - Future: Subsets can be performed based on a provider controlled access flag, which can be set according to quality concerns, spatial location, etc.

DMS State Machine



Data Management Key Components

- **Conditions**
  - Temporal
  - Rolling Temporal
  - Boolean
- **Rules**
  - Restrictions
  - Permissions
- **Groups**
  - Only used with Permissions

Conditions – What are They?

- **Phraselets (unable to be evaluated)**
  - Temporal
    - Start: Sept 1, 2002
    - Stop: Sept 31, 2002
  - Rolling Temporal
    - 30 days
    - 10 years
  - Boolean
    - True
    - False

Evaluating Conditions - Comparators

- **Comparators + Conditions = Evaluation Mechanism**
- **Examples of Comparators**
  - <, >, <=, >=, ==, !=
- **How they are used**
  - Comparator: ==
  - Temporal Condition: start Sept 1, 2002; stop Sept 30, 2002
  - Result: The time frame between September 1, 2002 and Sept 30, 2002
- **Questions**
  - How would this evaluated with a != comparator?
  - Does a < comparator make sense in this context?

Rules

- **Provider-Only access**
- **References**
  - Condition and Comparator
  - Provider
  - Granule or Collection
  - Group (optional)
- **Types of Rules**
  - Restriction (does not use a Group)
  - Permission (requires a Group)

Restrictions

- **A form of a rule**
- **Prevents access to metadata**
- **Applies to all ECHO users**
- **Specific to an Action**
  - Actions: View or Order
- **Specific to a piece of metadata**
  - Collection MOD01
  - Granule SC:1234

Permissions

- **A form of a rule**
- **Enables access to metadata**
- **Applies to a group previously created**
- **Specific to an Action**
  - Actions: View or Order
- **Specific to a piece of metadata**
  - Collection MOD01
  - Granule SC:1234

Restriction Example

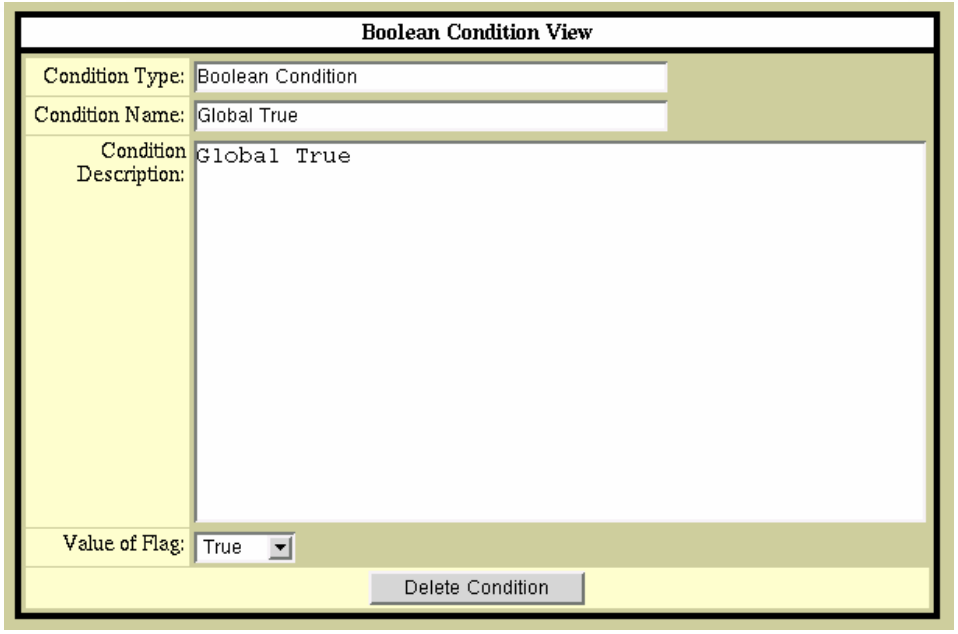
- **Desired Behavior: A provider wishes to prevent viewing of all "young" collections.**
- **Condition Information**
  - ConditionType: Rolling Temporal
  - Duration: 30 Days
- **Rule Information**
  - RuleType: Restriction
  - Comparator: Less Than
  - DataType: Collection
  - DataValue: "ALL"

Permission Example

- **Desired Behavior: A provider wishes to grant viewing access to their internal testers**
- **Condition Information**
  - Condition Type: Boolean
  - Value: True
- **Rule Information**
  - RuleType: Permission
  - Comparator: Equals
  - DataType: Collection
  - DataValue: “ALL”
  - Group: Internal Testers

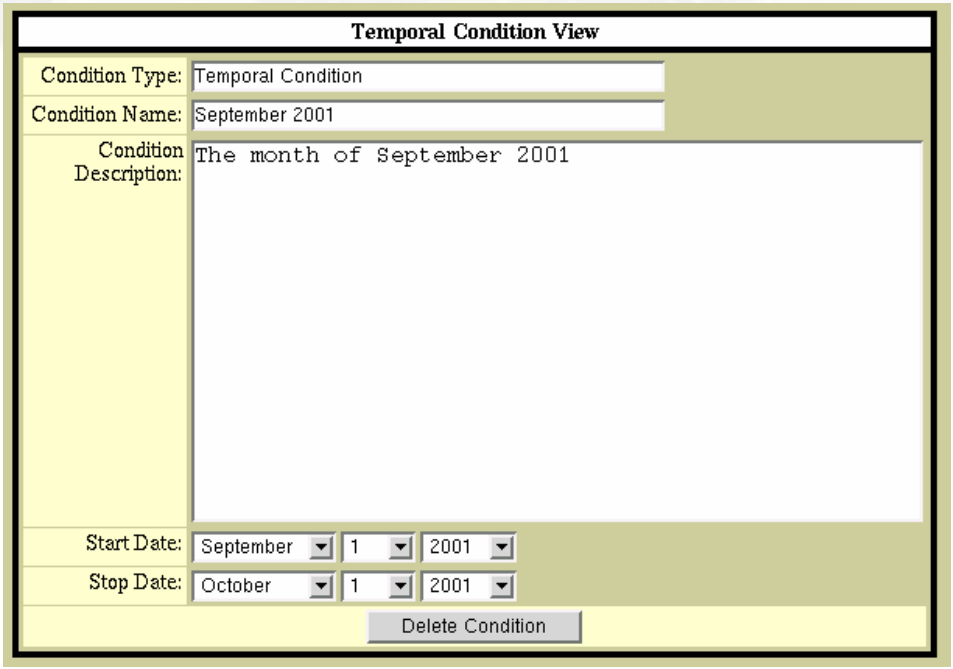
Data Management Service
(Conditions)

- **Condition Management is separate from Data Rule Management**
- **Shows how to create a globally true condition.**



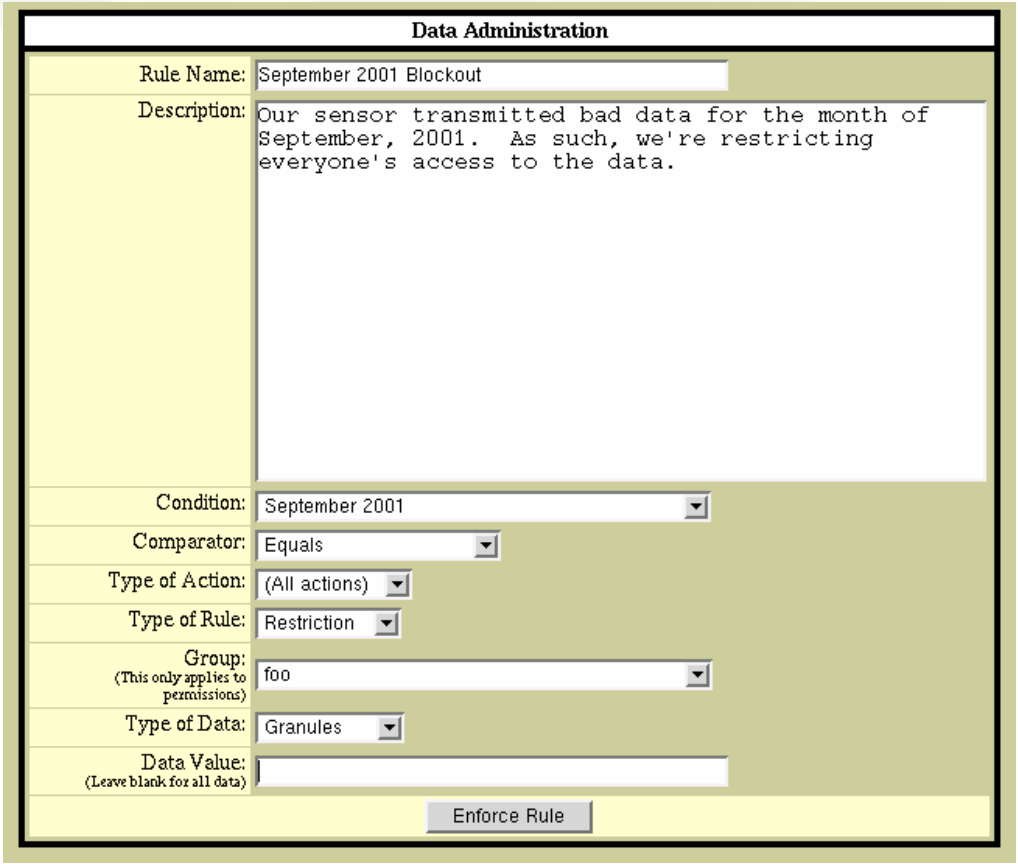
Data Management Service
(Conditions)

- **Shows how to create a condition based on the temporal time frame of September, 2001**



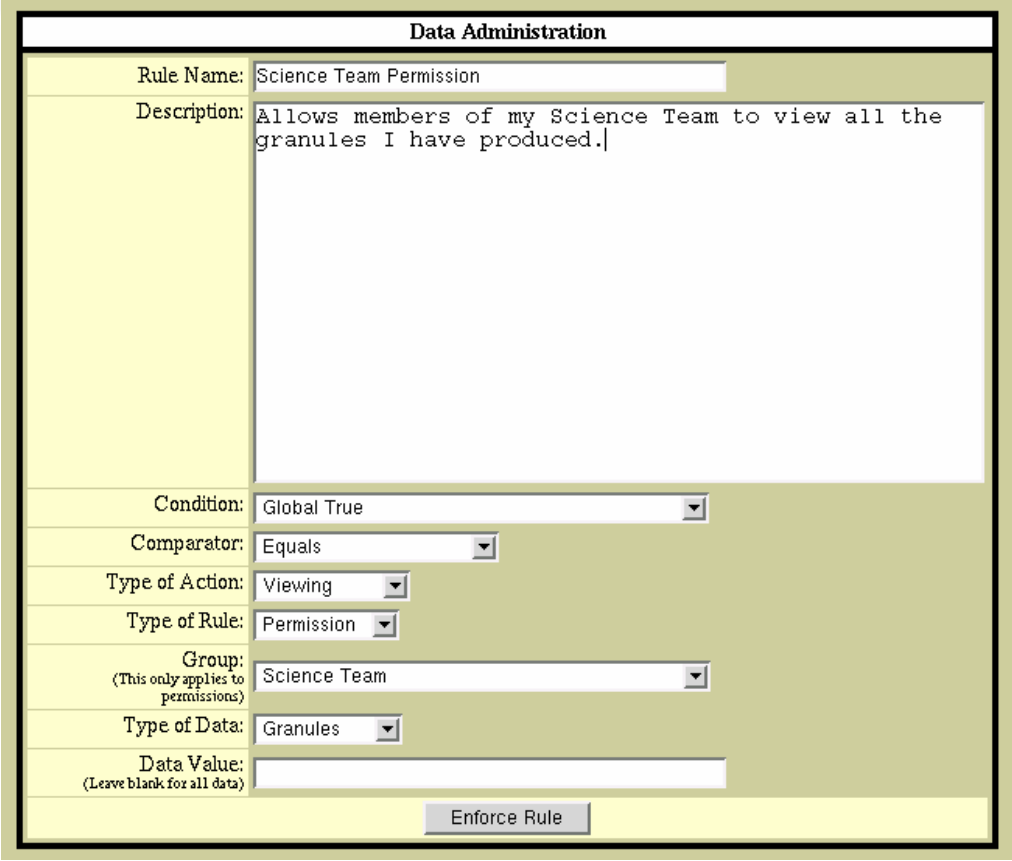
Data Management Service
(Restrictions)

- Restrictions apply to all users.
- Restrictions are evaluated using a Condition and Comparator
- Restrictions can apply to a particular granule, or all granules (if blank)



Data Management Service
(Permissions)

- **Permissions apply to a particular group**
- **Permissions can only use Boolean conditions**
- **Permissions can apply to a particular granule, or all granules**



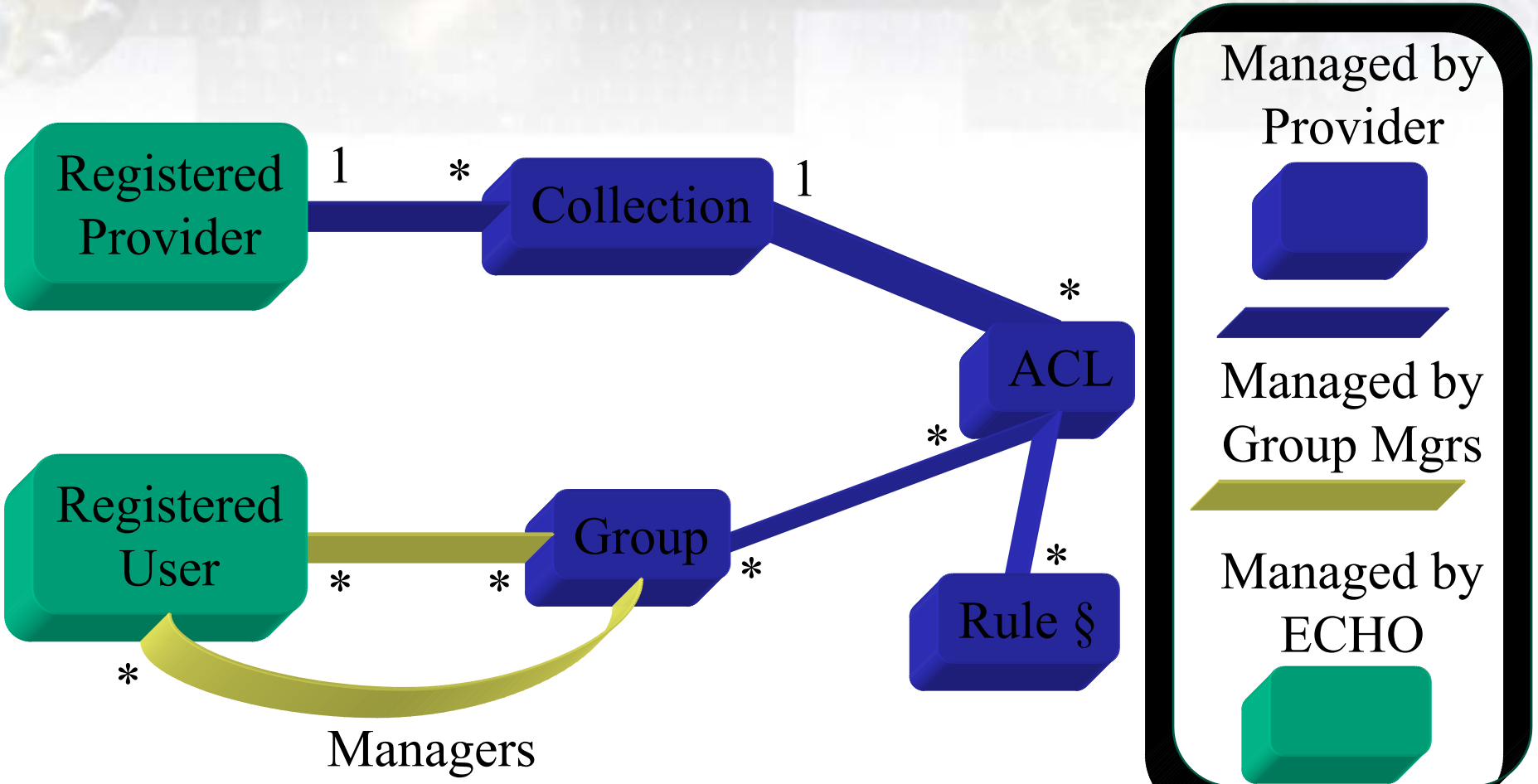
Visibility “Gotcha’s”

- **Rules evaluated optimistically**
  - Permissions supercede Restrictions
- **Rules short-circuit**
  - “ALL” DataValue keyword short-circuits individual DataValues
  - As soon as permission is located, visibility is granted
- **Conditions are re-usable**
- **Group management an external process**
- **It is possible using both the GUI and API to create ACLs that make no sense**

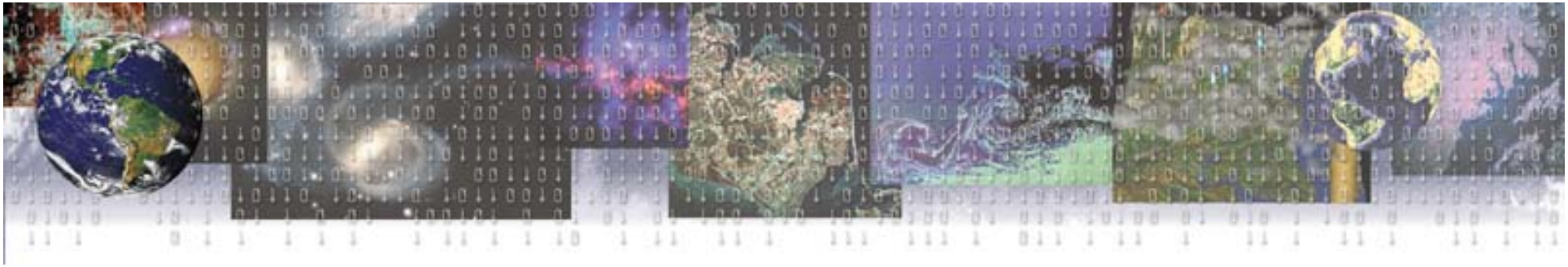
5.0 Improvements

- **Providers are no longer users who log in**
  - Registered users have provider role
- **Management of groups is now limited to registered users**
- **New easier to use user interface**

Metadata Visibility – Conceptual Model



§Rule determines if collection is hidden and if not, which granules are hidden



PUMP

ECHO

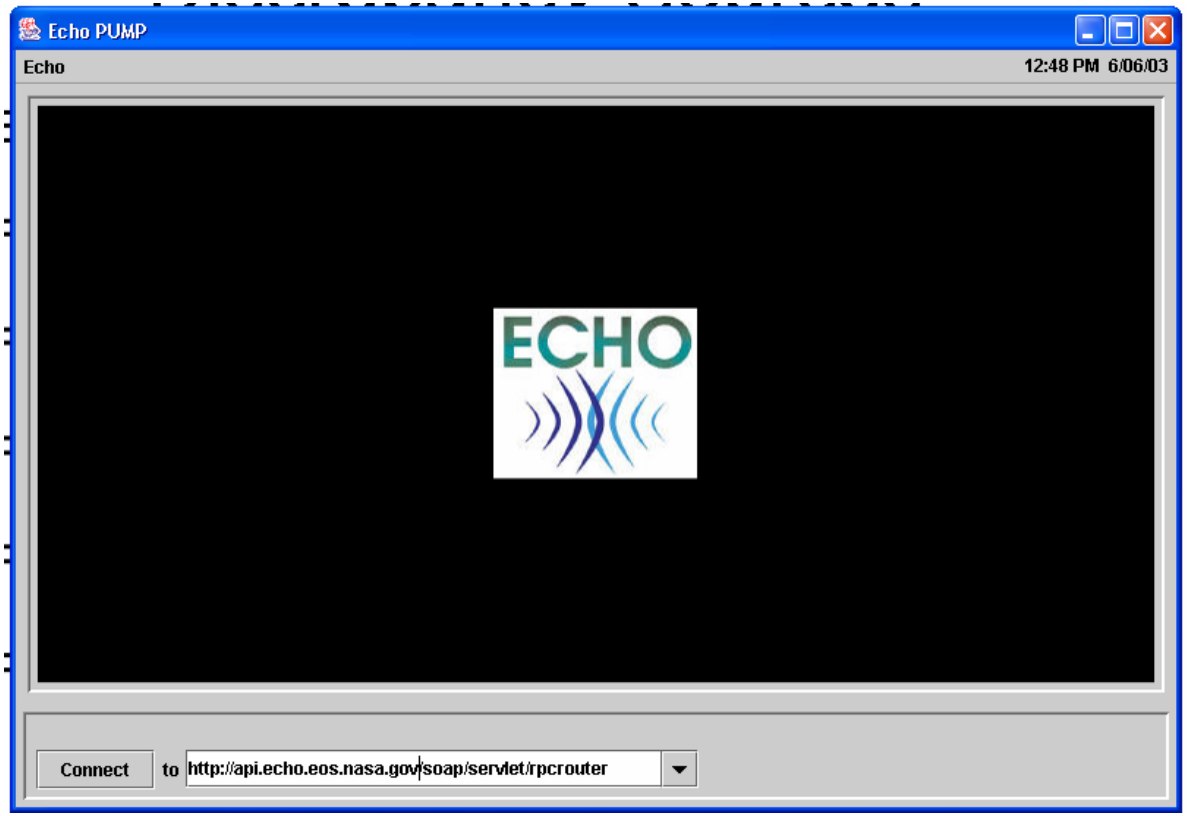
Global Science & Technology, Inc.

PUMP Attributes

- **PUMP – The Provider User Management Program**
- **Java application**
  - Installs with install anywhere COTS product
- **Addresses providers and administrators needs**
  - Account Management
  - User Management
  - ACL Management
  - More to come…

PUMP Connection

- **Specify which instance of ECHO to connect to**
  - Allows connection to "test" instances of ECHO

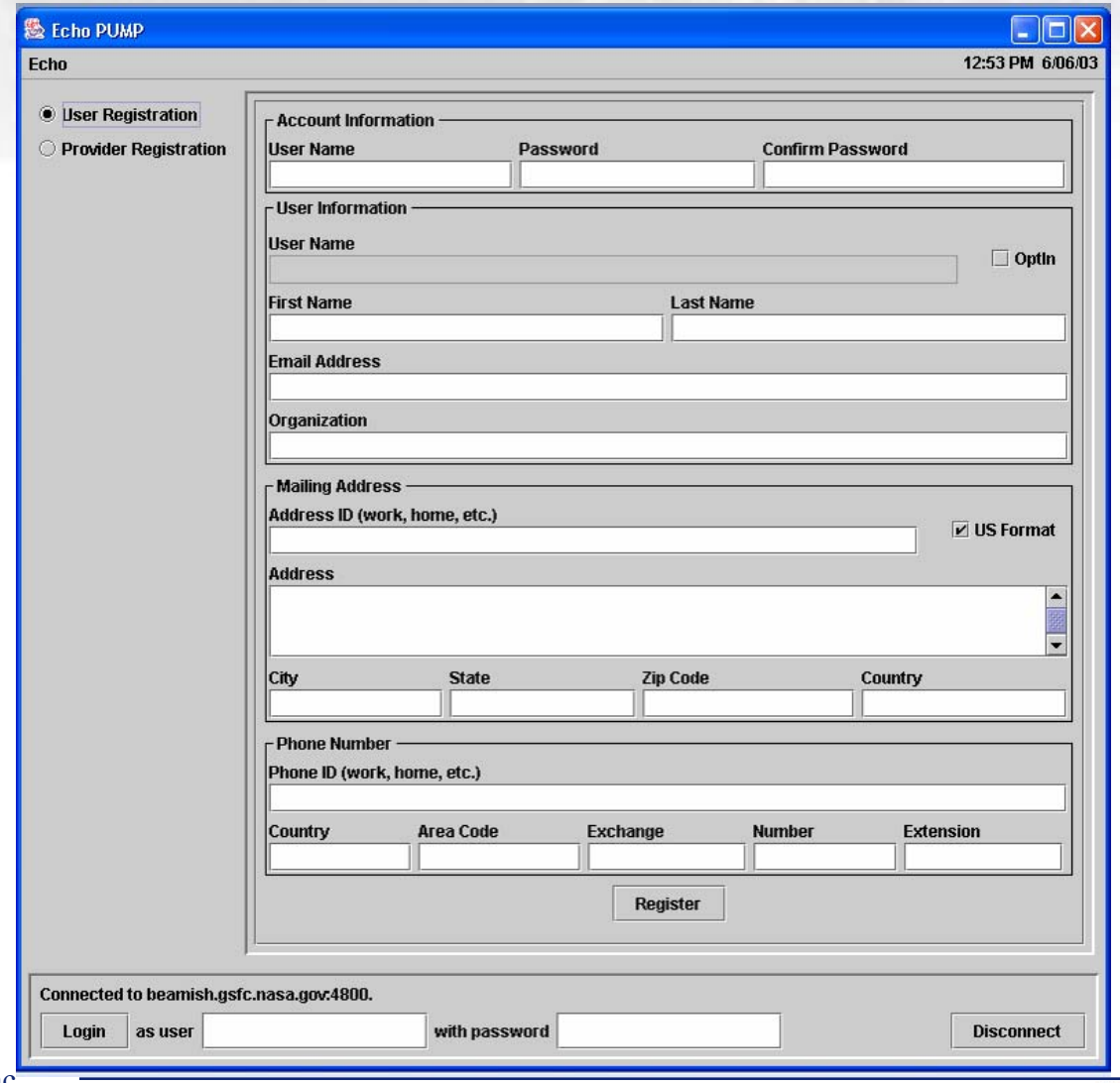


PUMP – First Step – Guest Context

- **After connecting to an instance of ECHO, the user can:**
  - Register a user
  - Register a provider
  - Login
  - Disconnect
- **Pane on left allows user to select functions**
- **Pane on bottom monitors connection and context**
- **Main pane shows panels customized to the selected function**

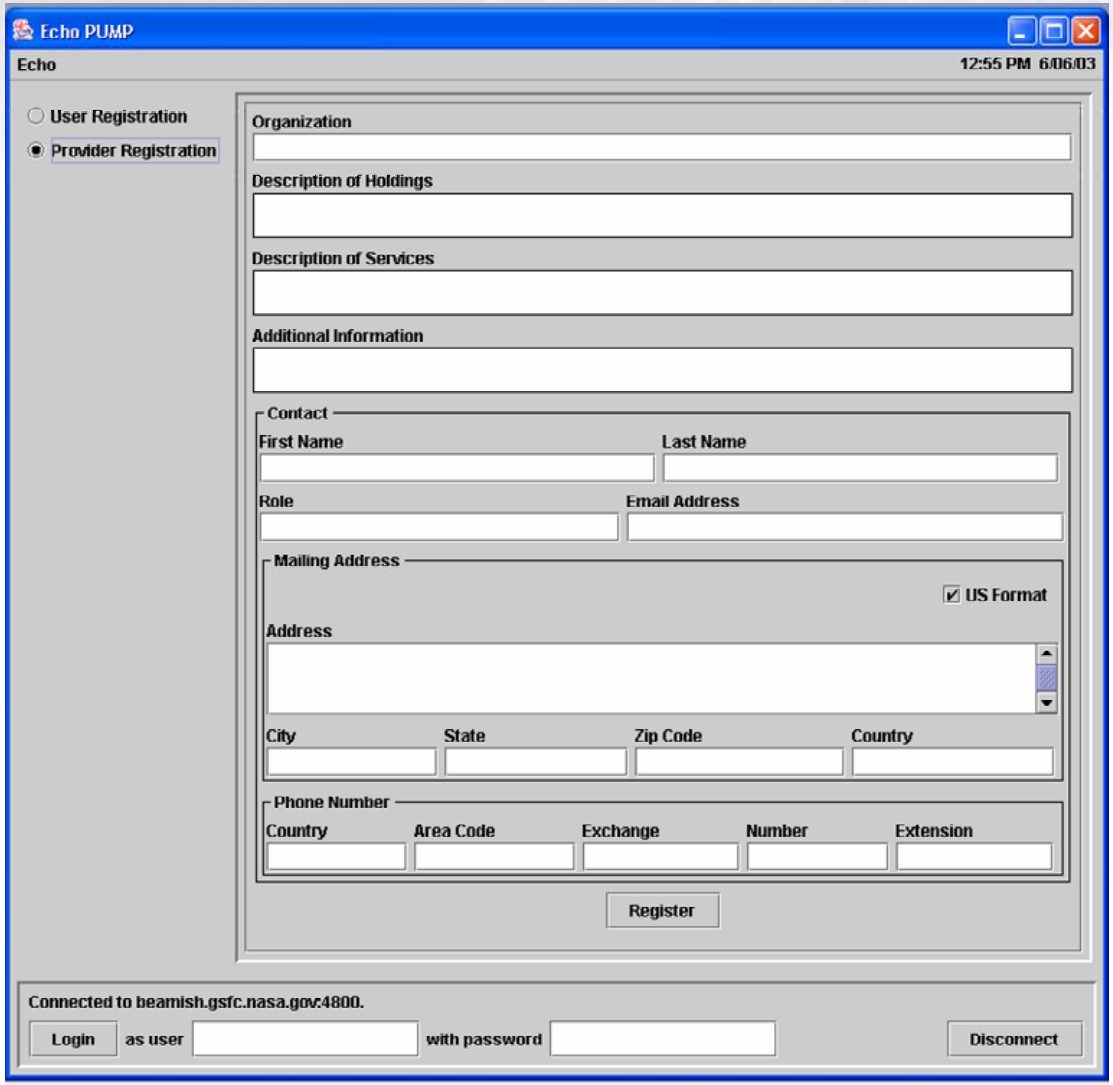
User Registration

- **Creates user with initial mailing address and phone number**
  - Others can be added later
  - All fields are required by PUMP except extension



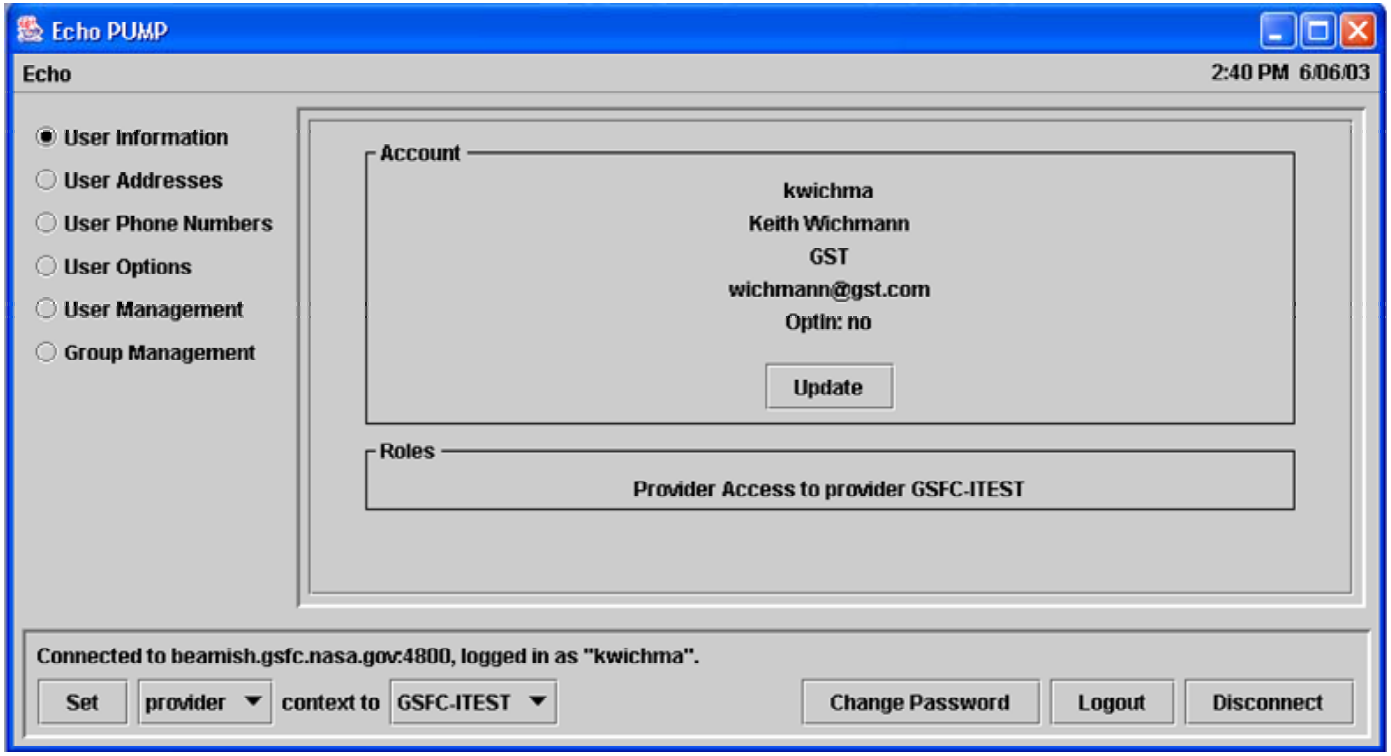
Provider Application Registration

- When a new provider wants to participate in ECHO, this form is used to indicate their holdings and contact information

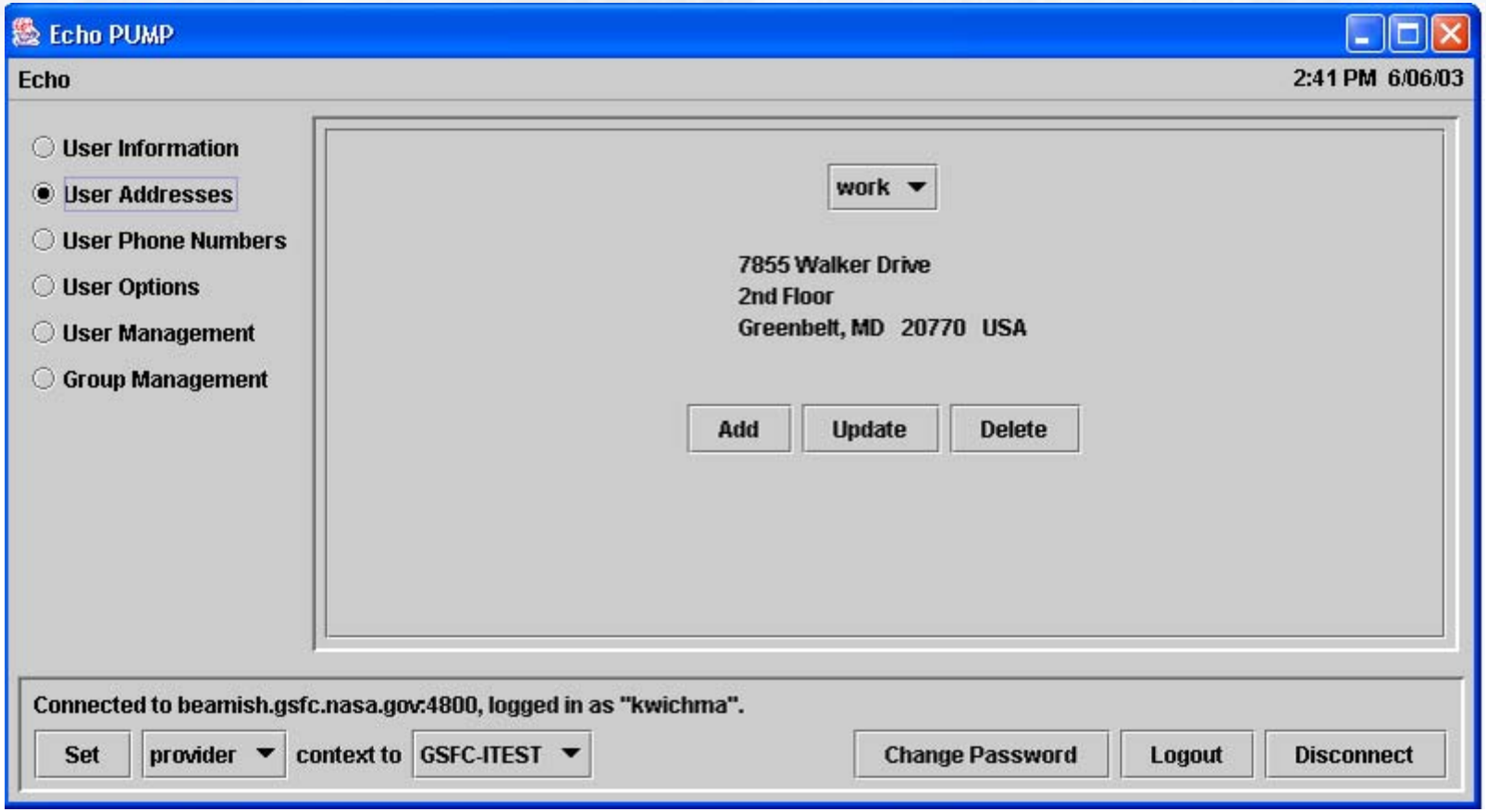


Main User Context

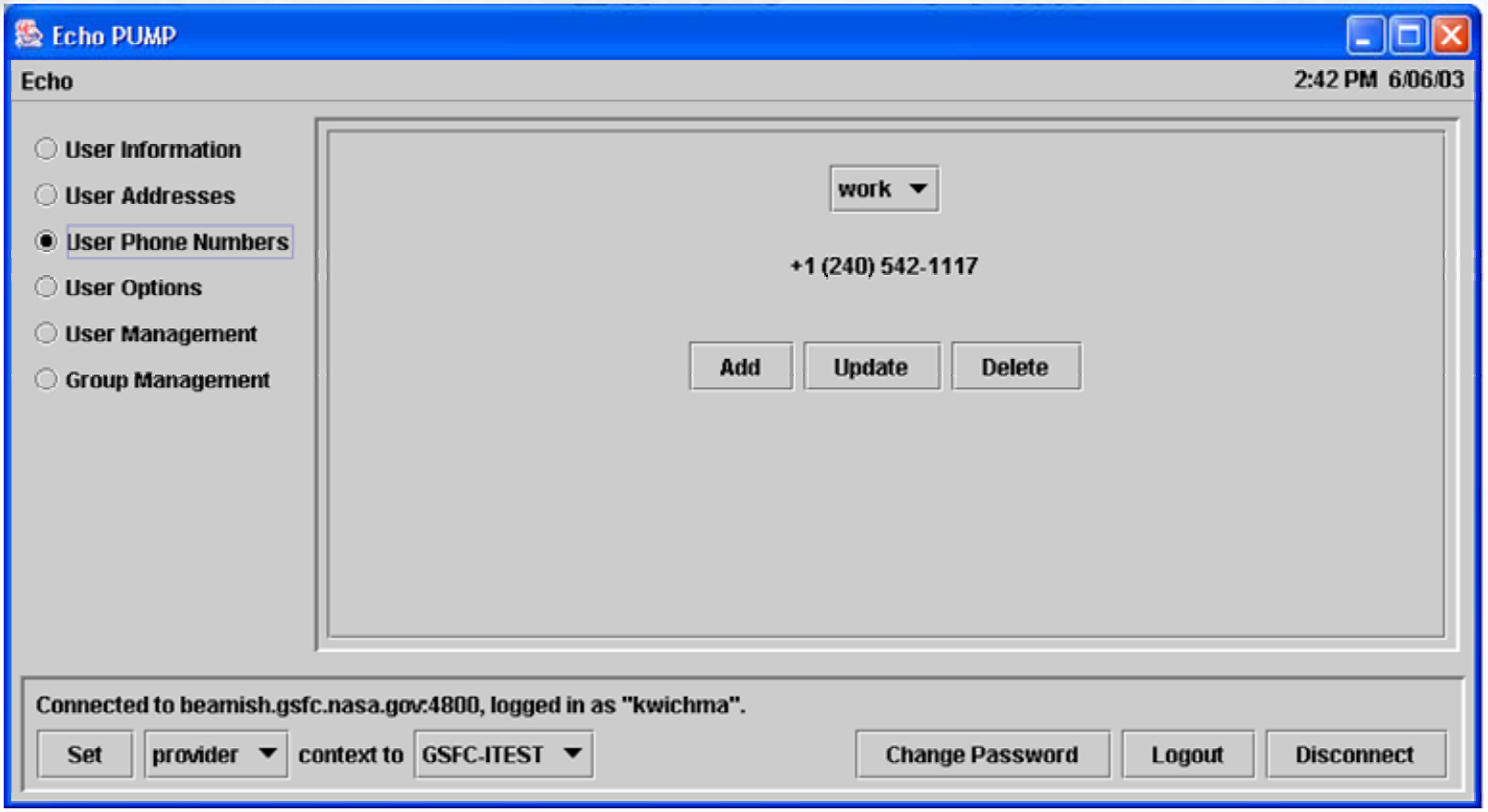
- These are the functions available to all registered users
- User roles are listed
- Ability to update name, organization, opt-in, and email



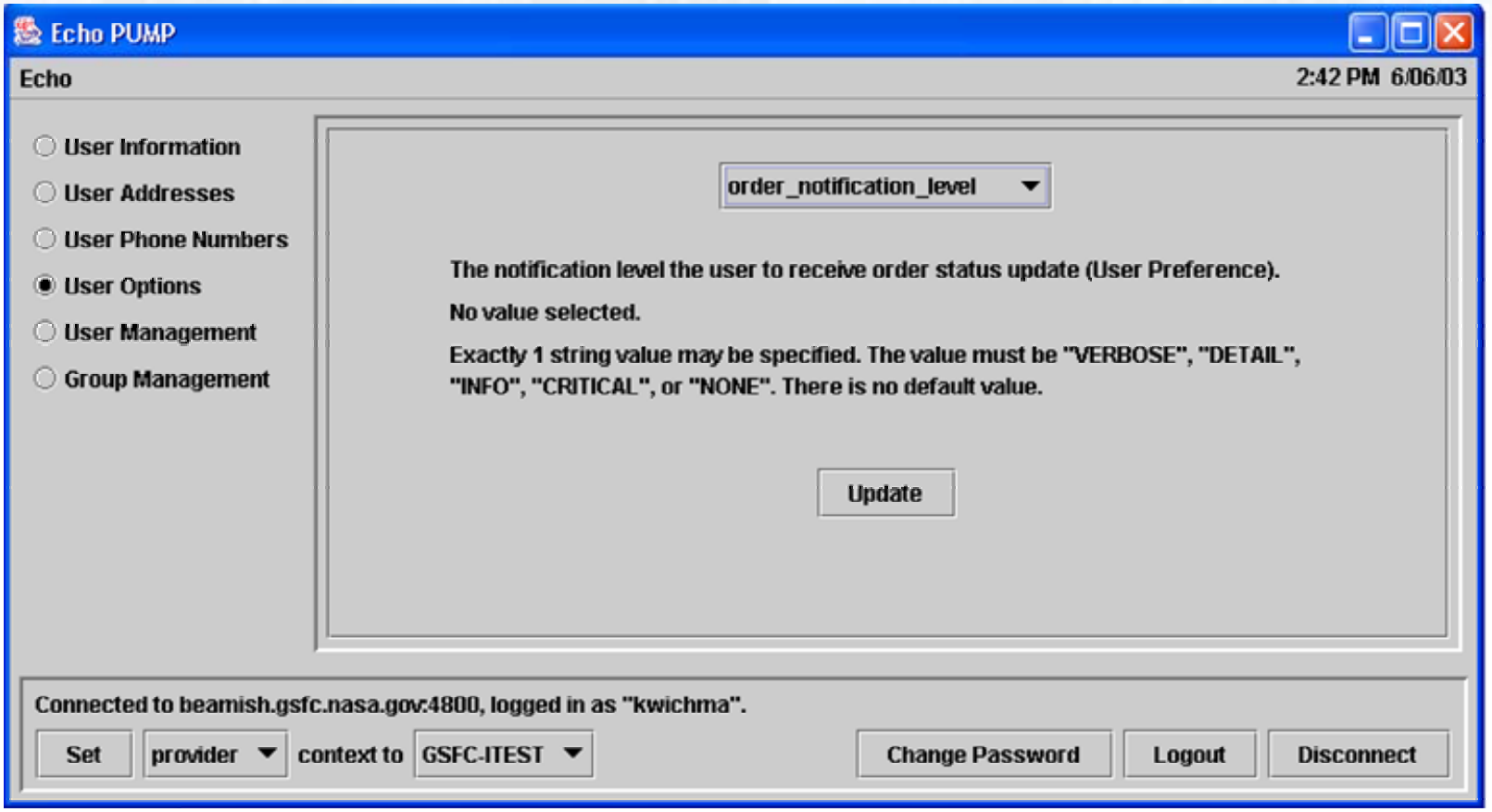
Manage User Addresses



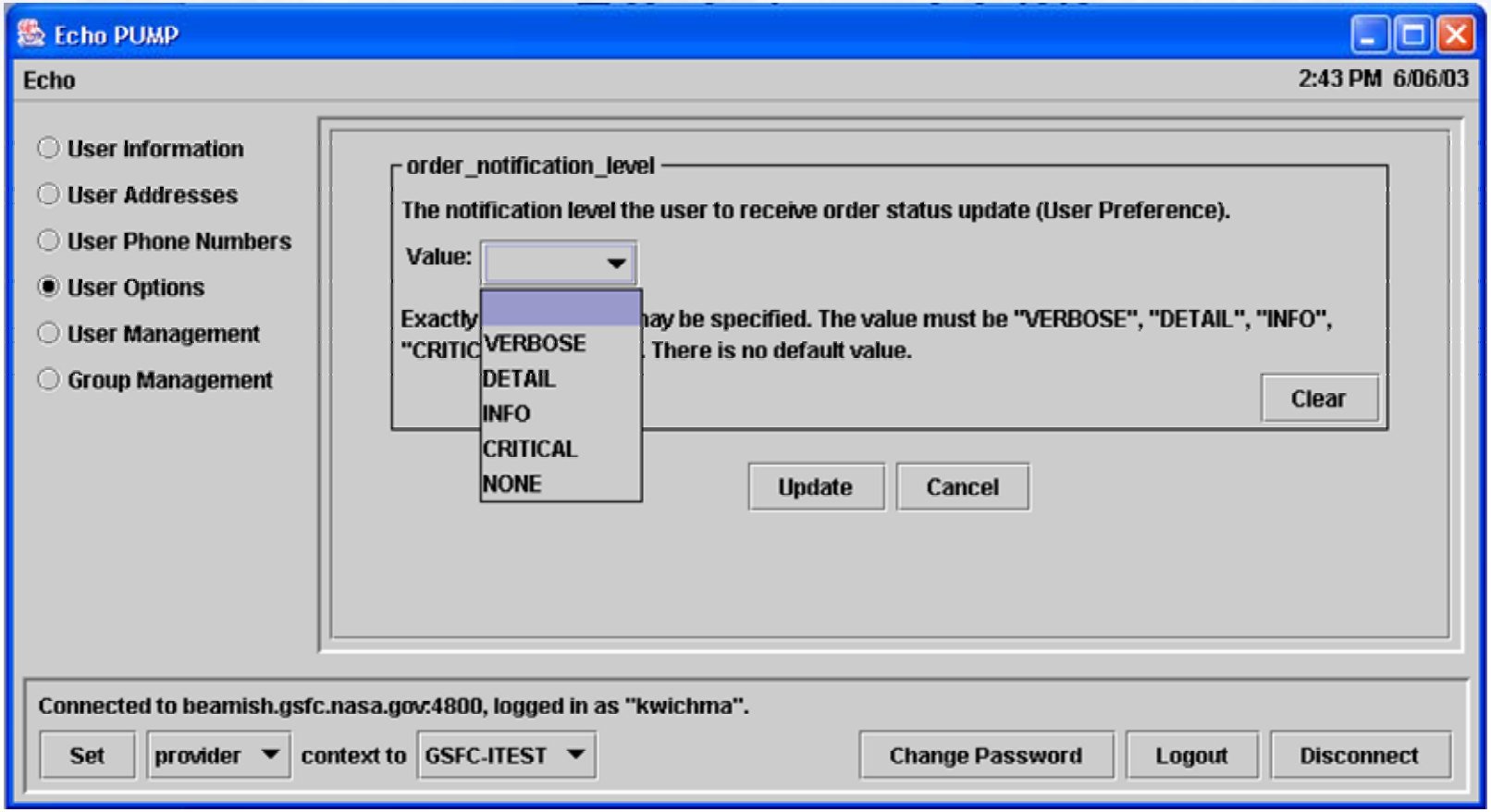
Manage User Phone Numbers



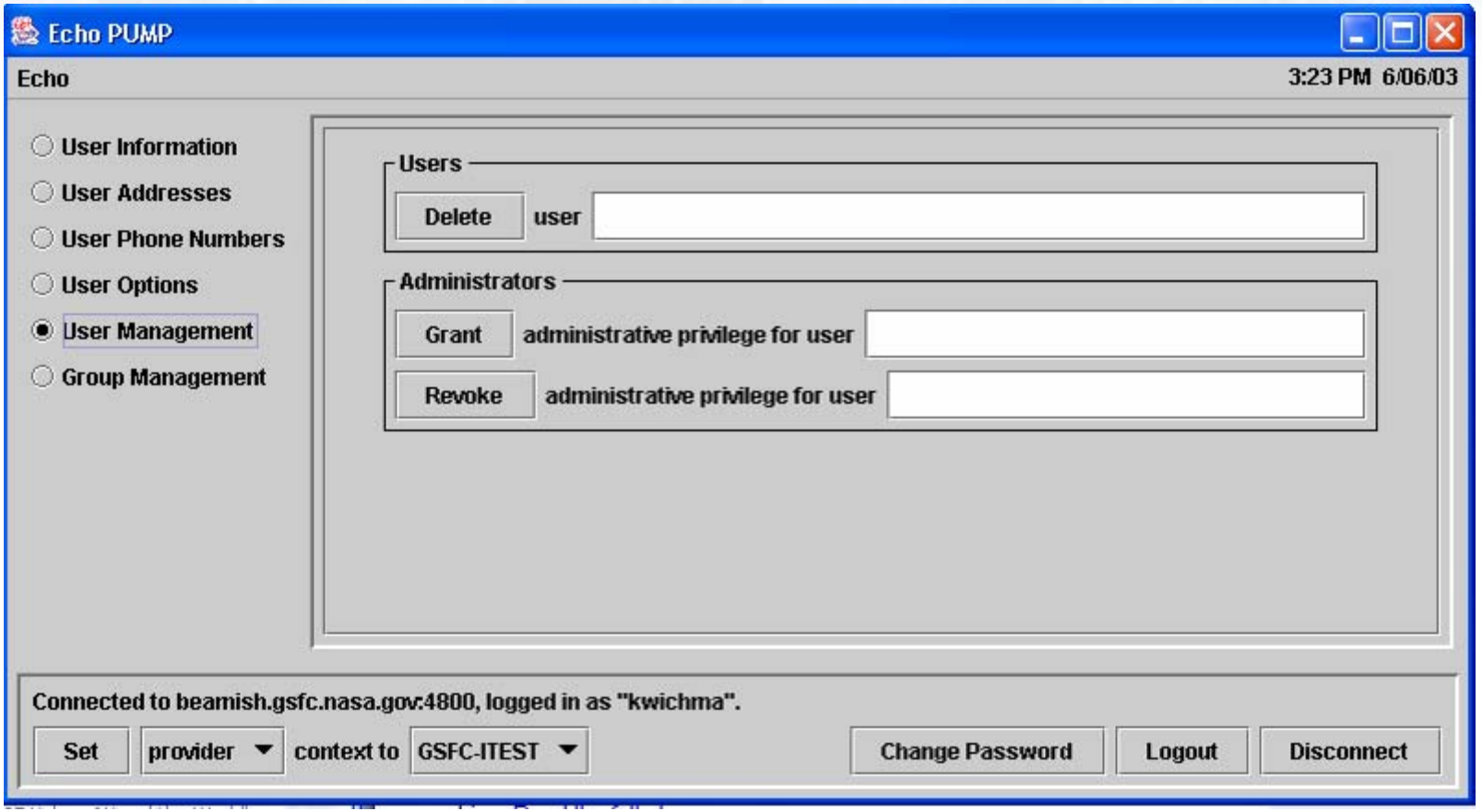
Manage User Options



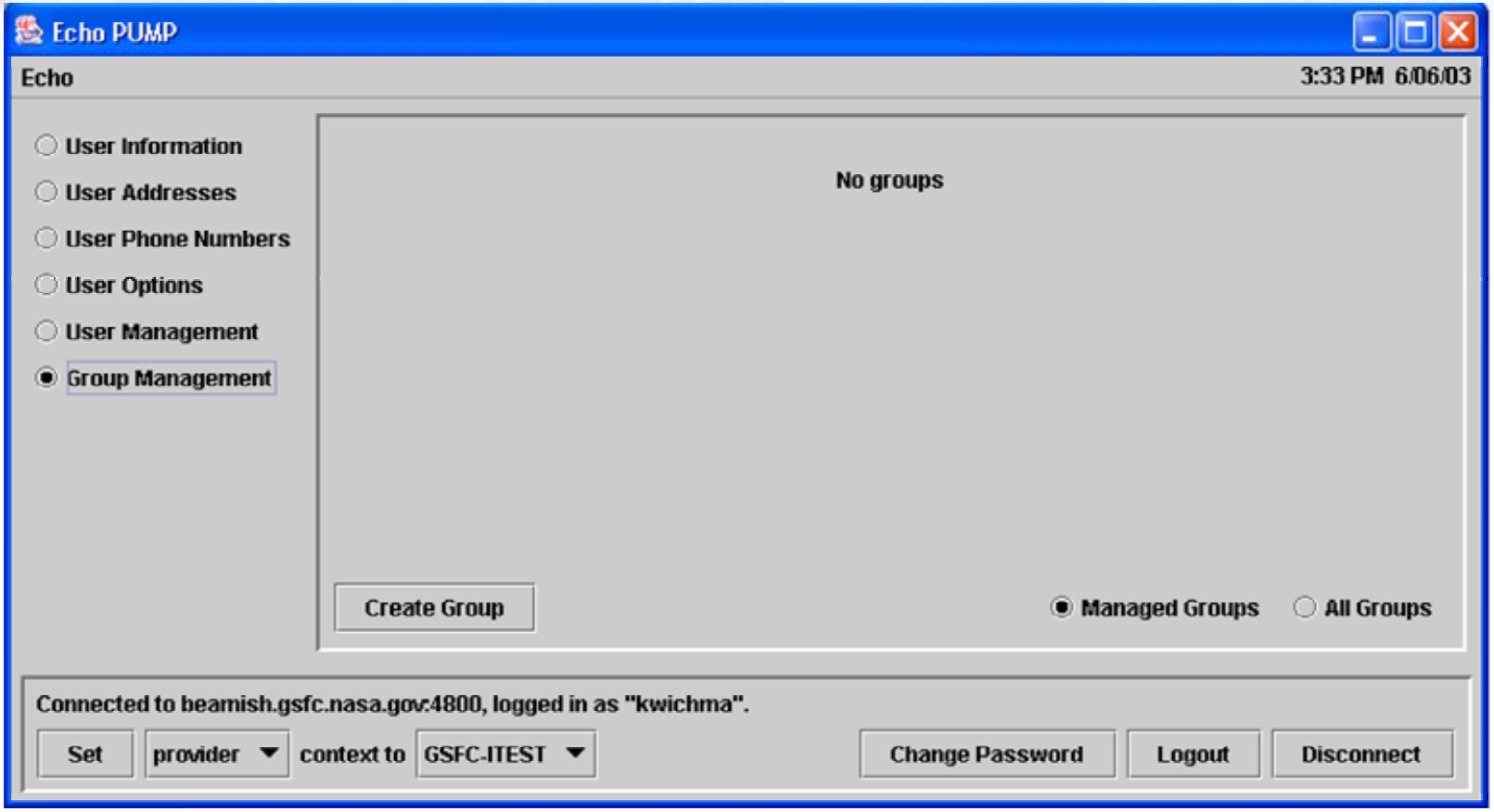
Set User Options



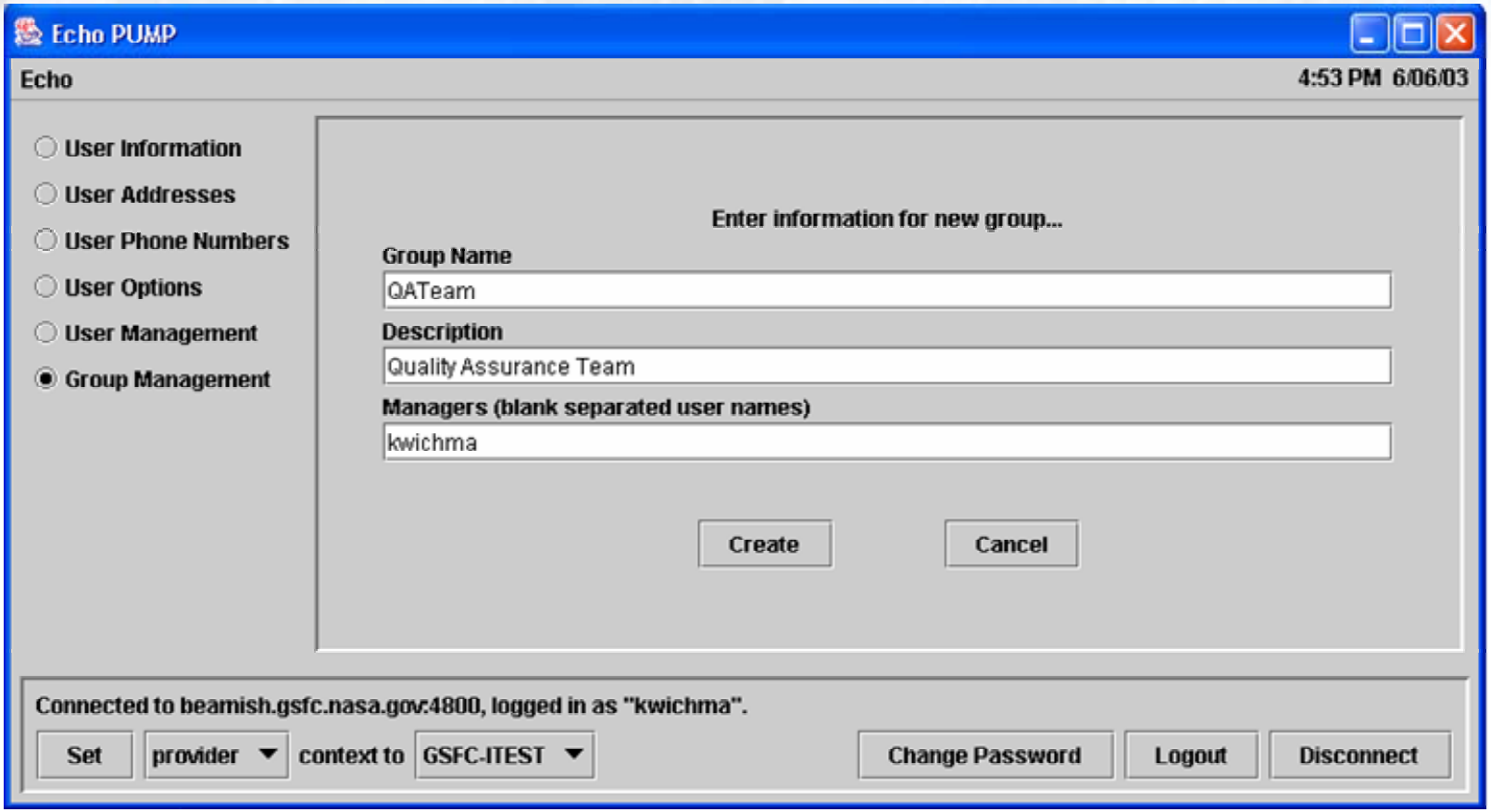
Administrative Functions – Admin Context



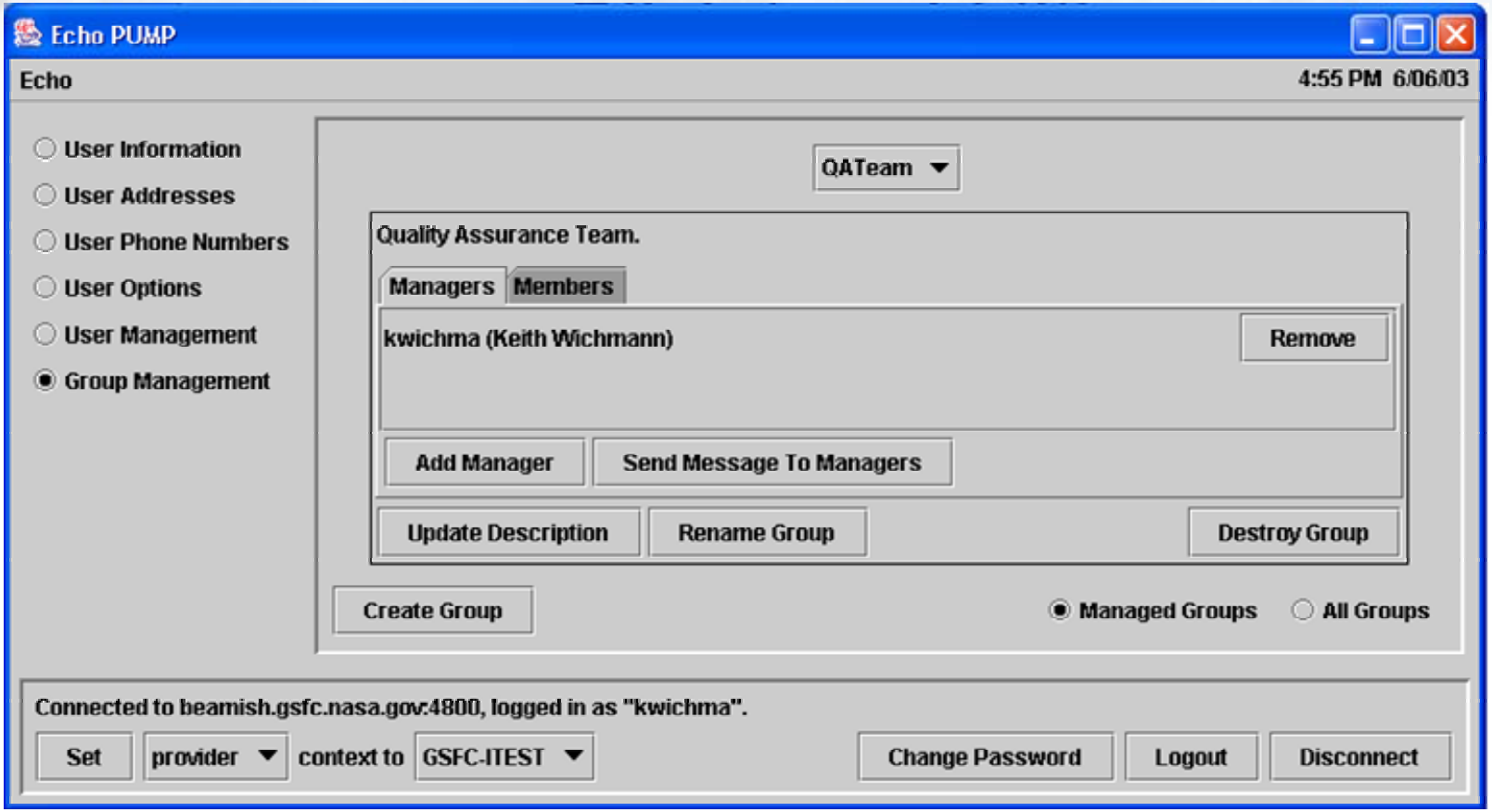
Group Management Functions



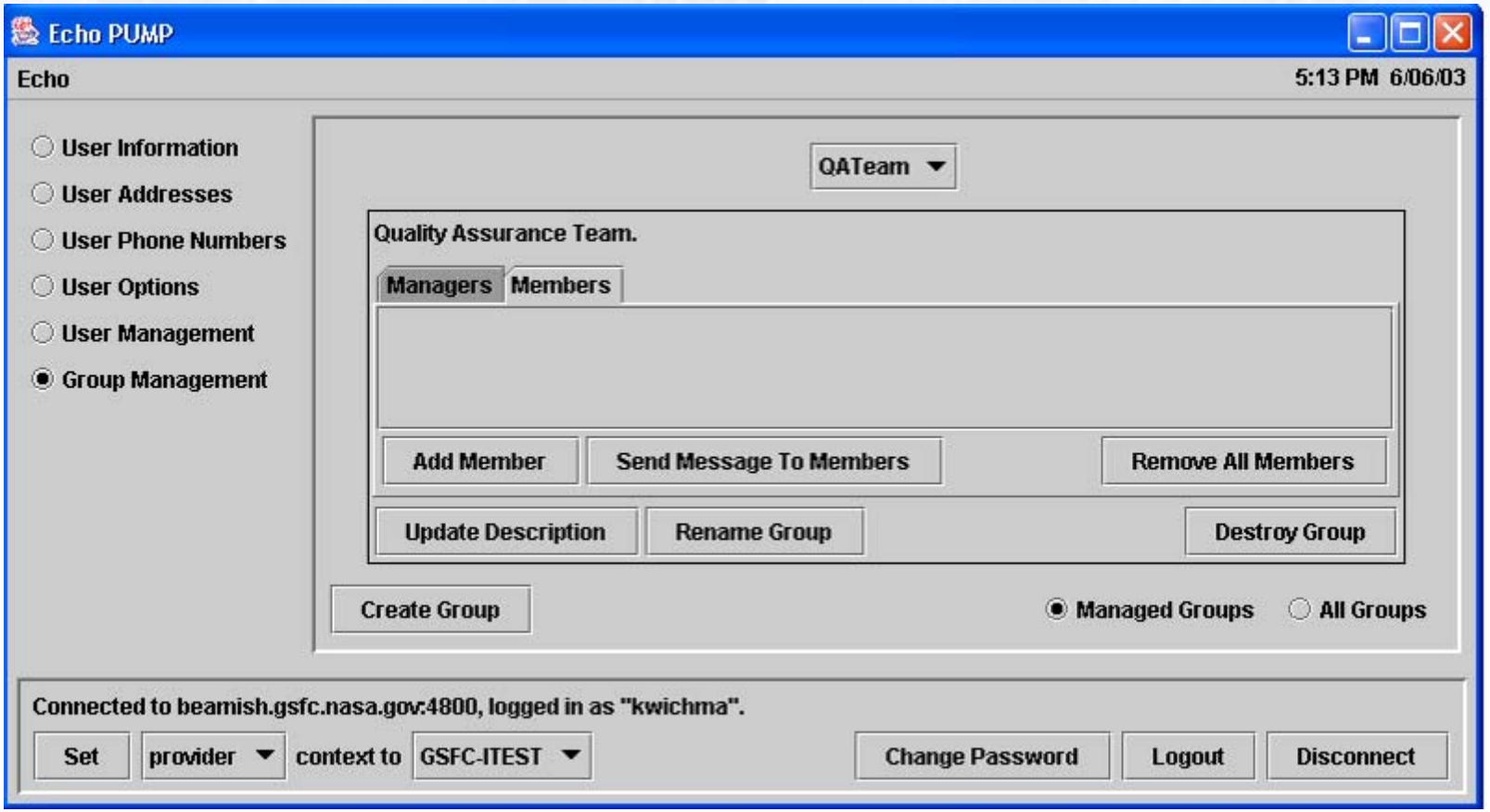
Create Group



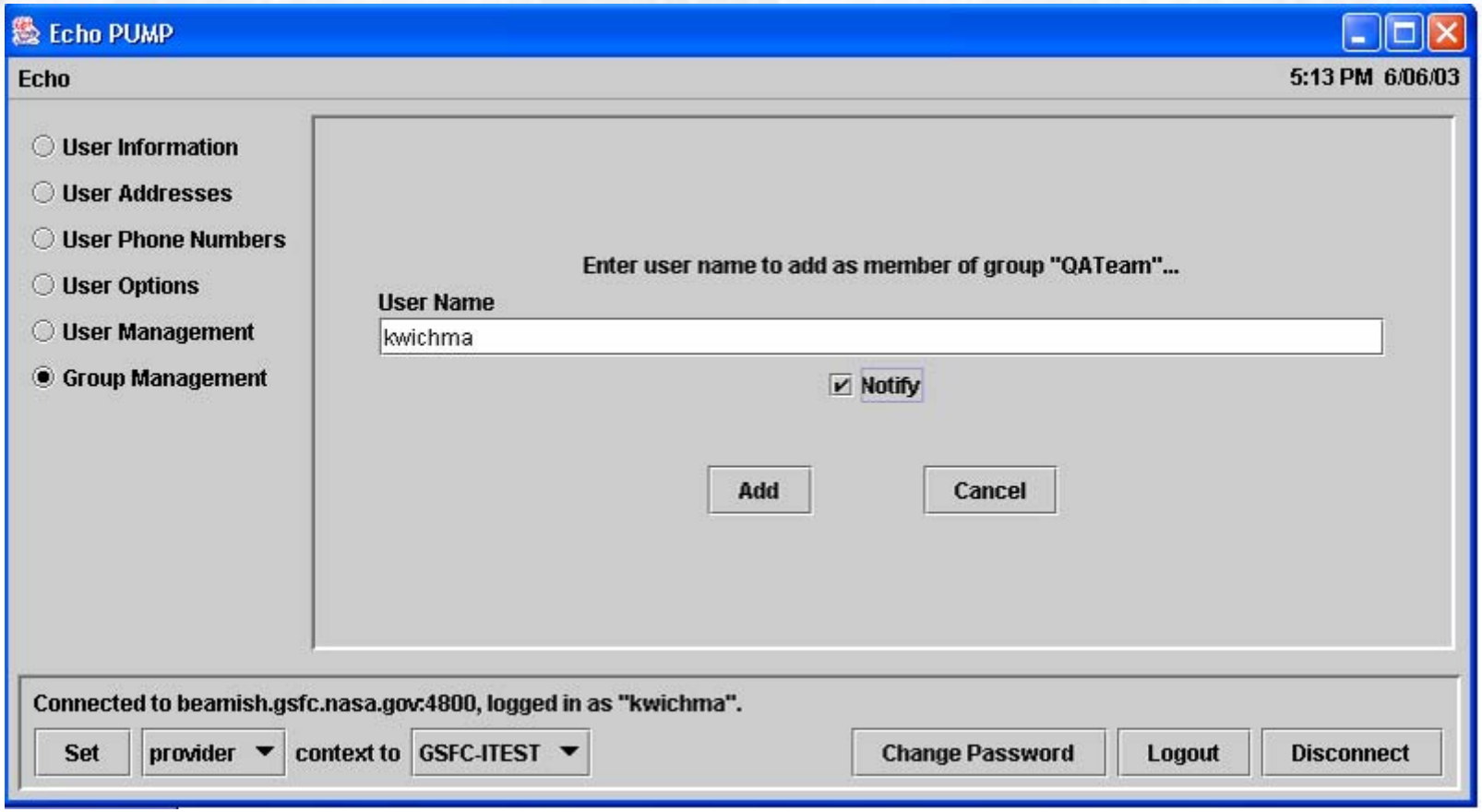
Group Display - Managers



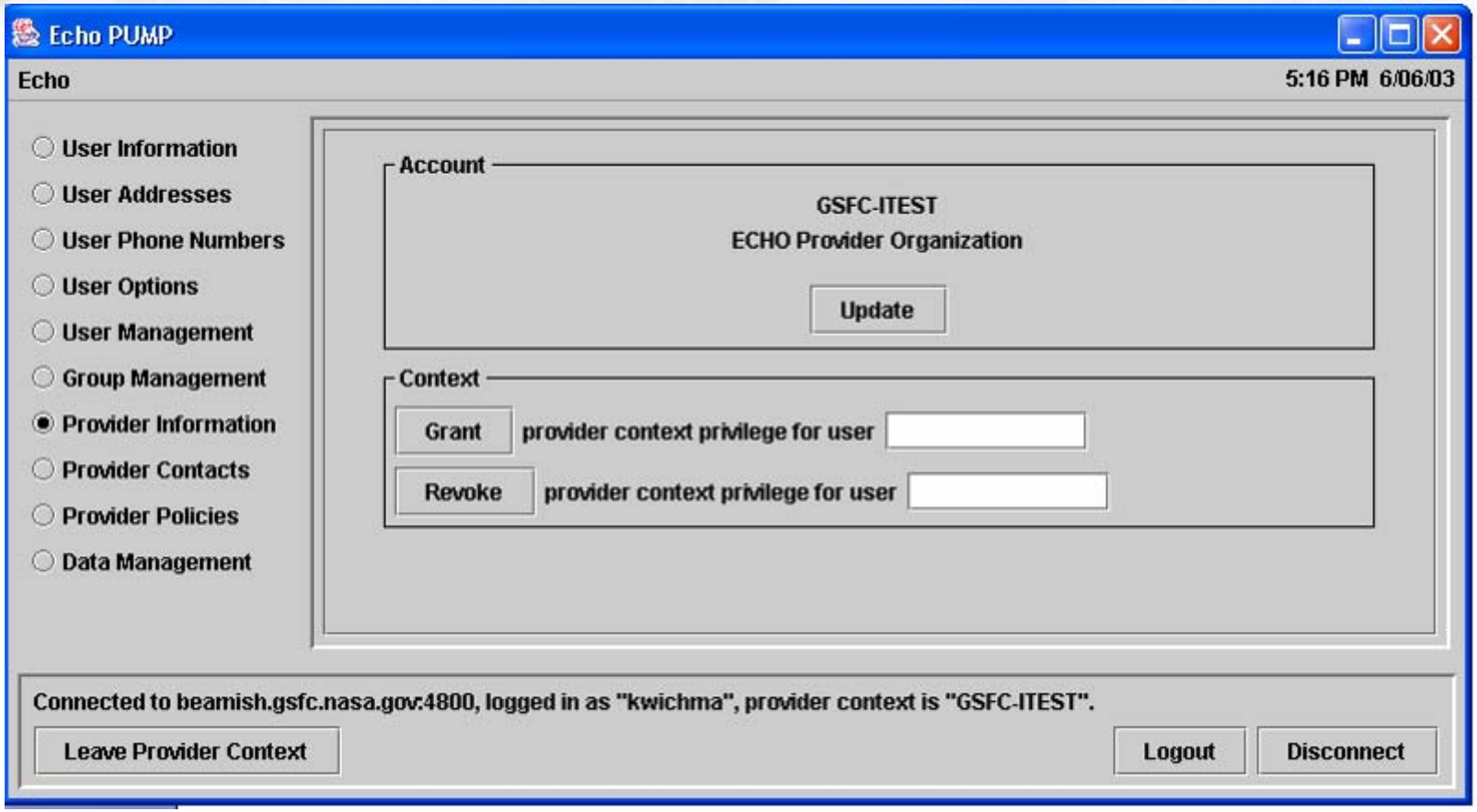
Group Display - Members



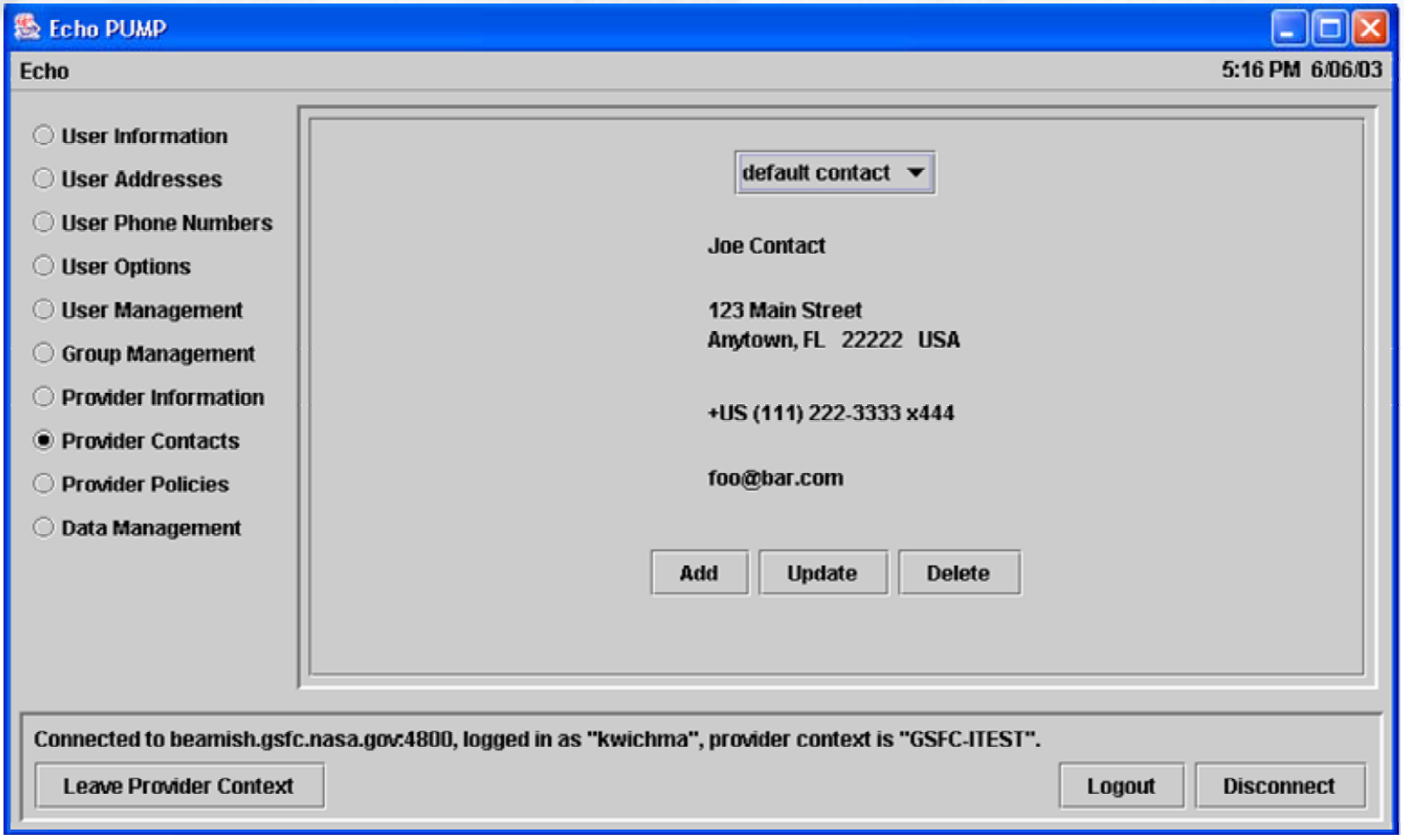
Group Management – Add Member



Provider Context – Account Description and Roles



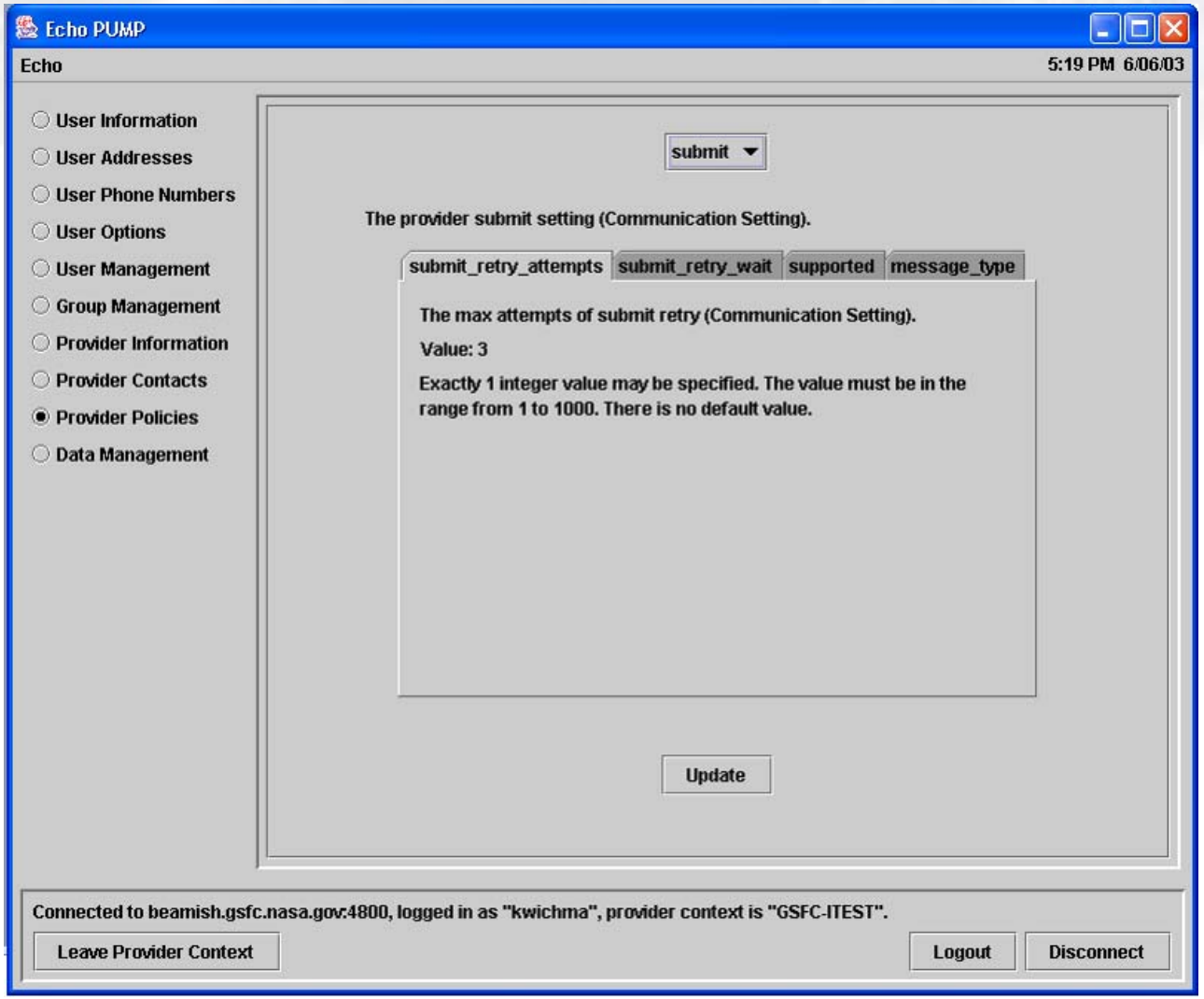
Manage Provider Contacts – Provider Context



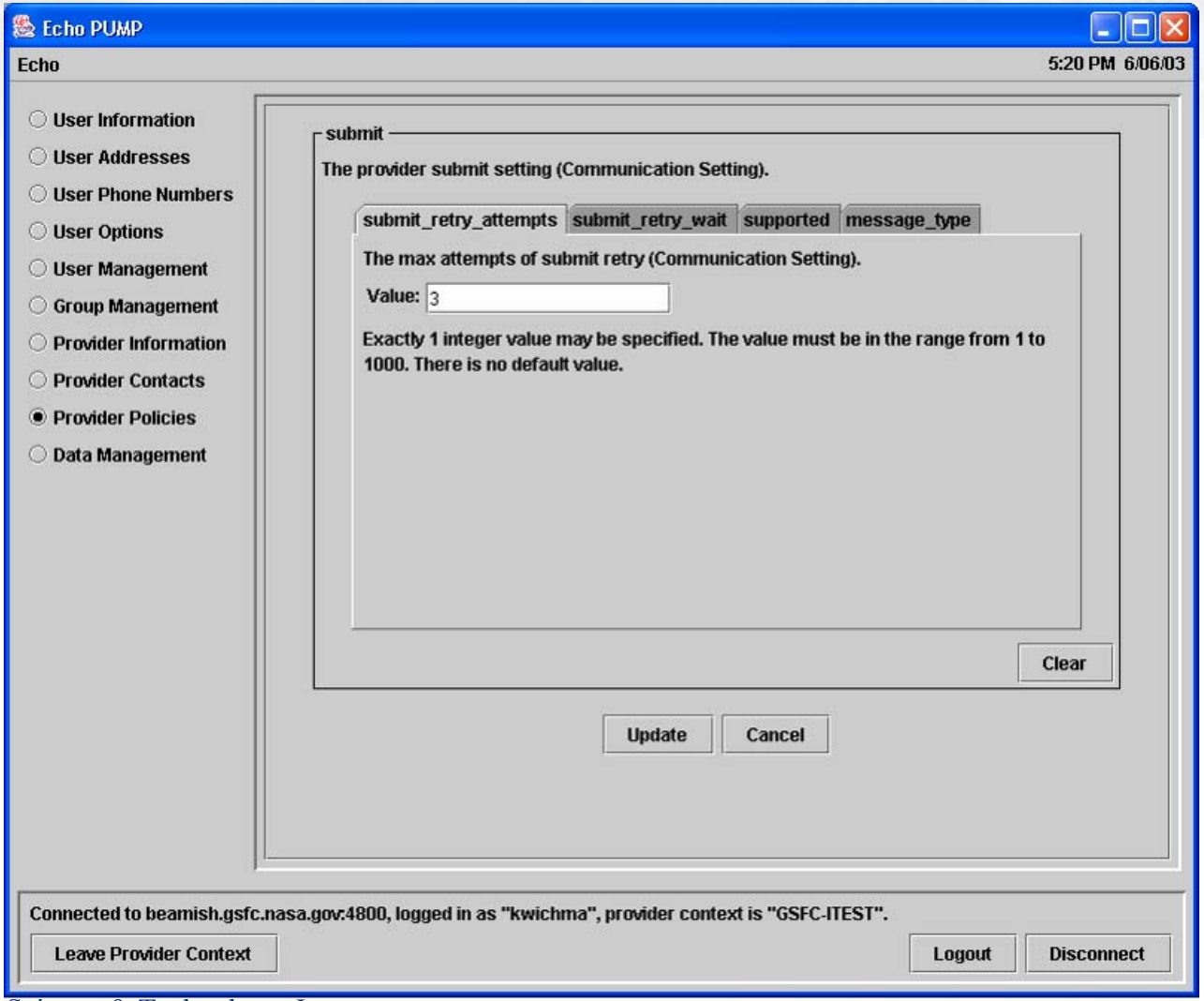
Provider Policies

- **For each of submit, quote, cancel**
  - Is it supported?
  - What message format?
    - ODL via sockets vs. XML via SOAP
  - Where to send the message?
  - Number of times to retry message send
  - Number of seconds to wait between retries

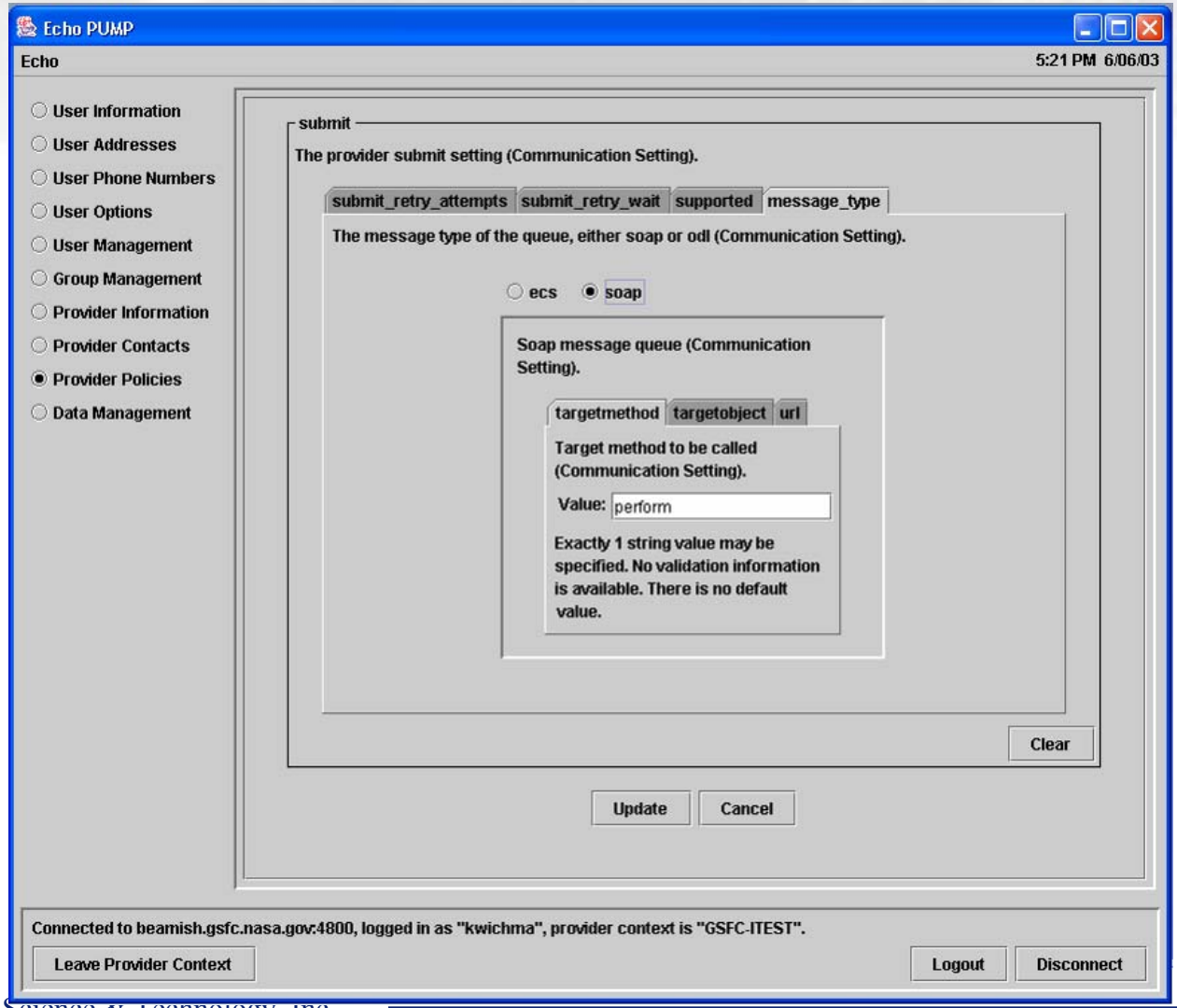
Provider Policies – Provider Context



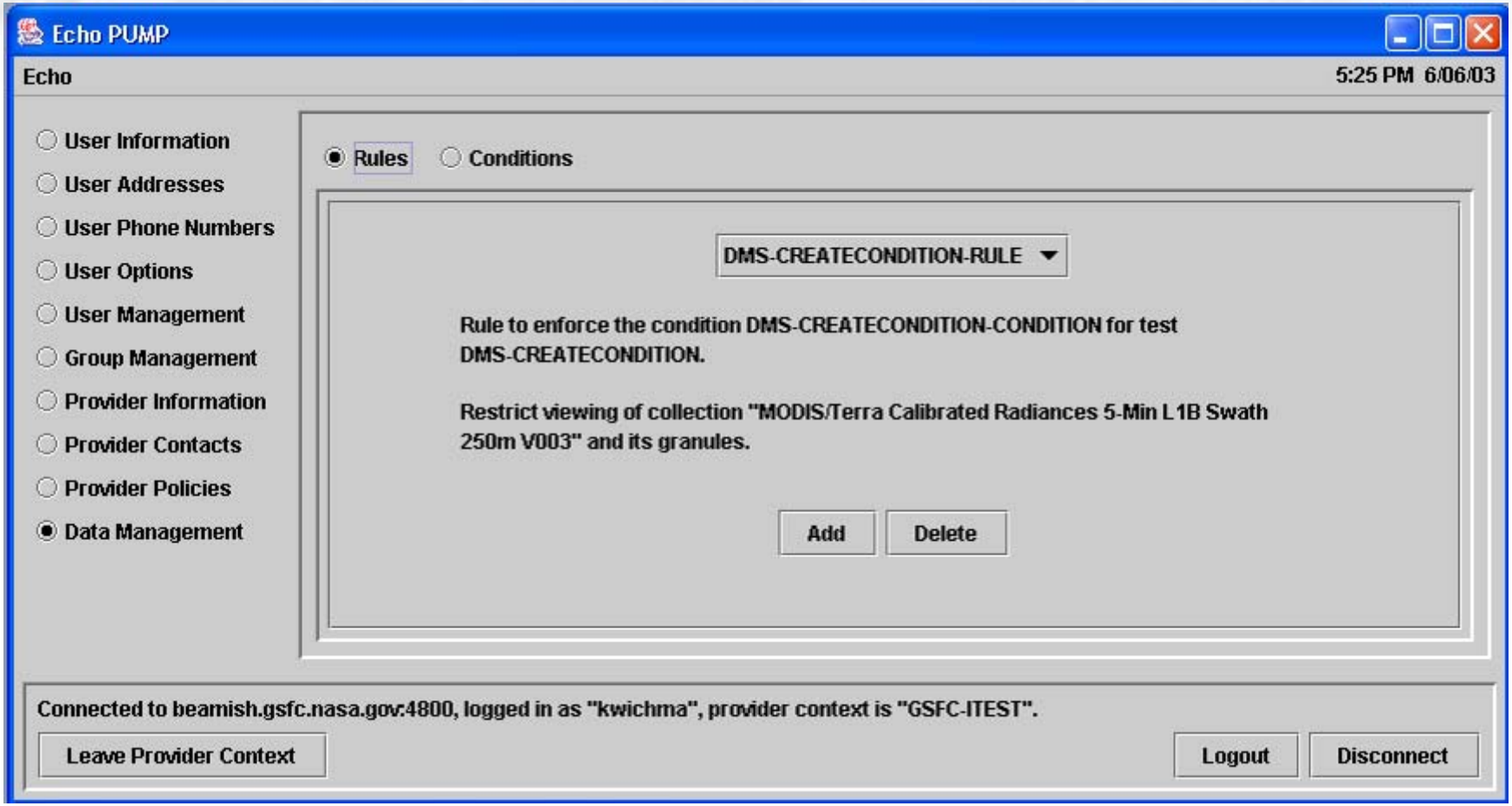
Set Provider Policies - Retries – Provider Context



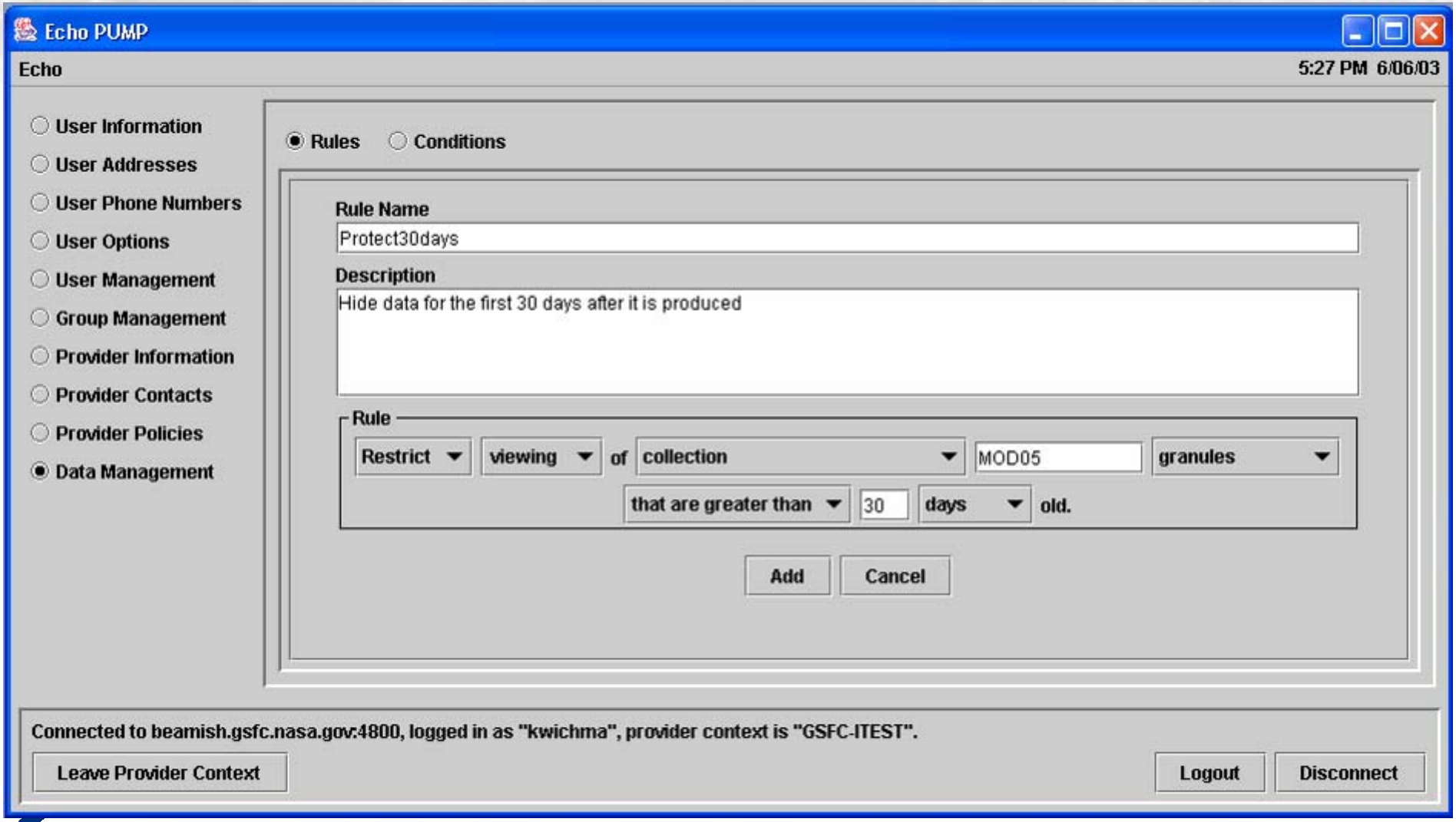
Set Provider Policies – Message Type – Provider Context



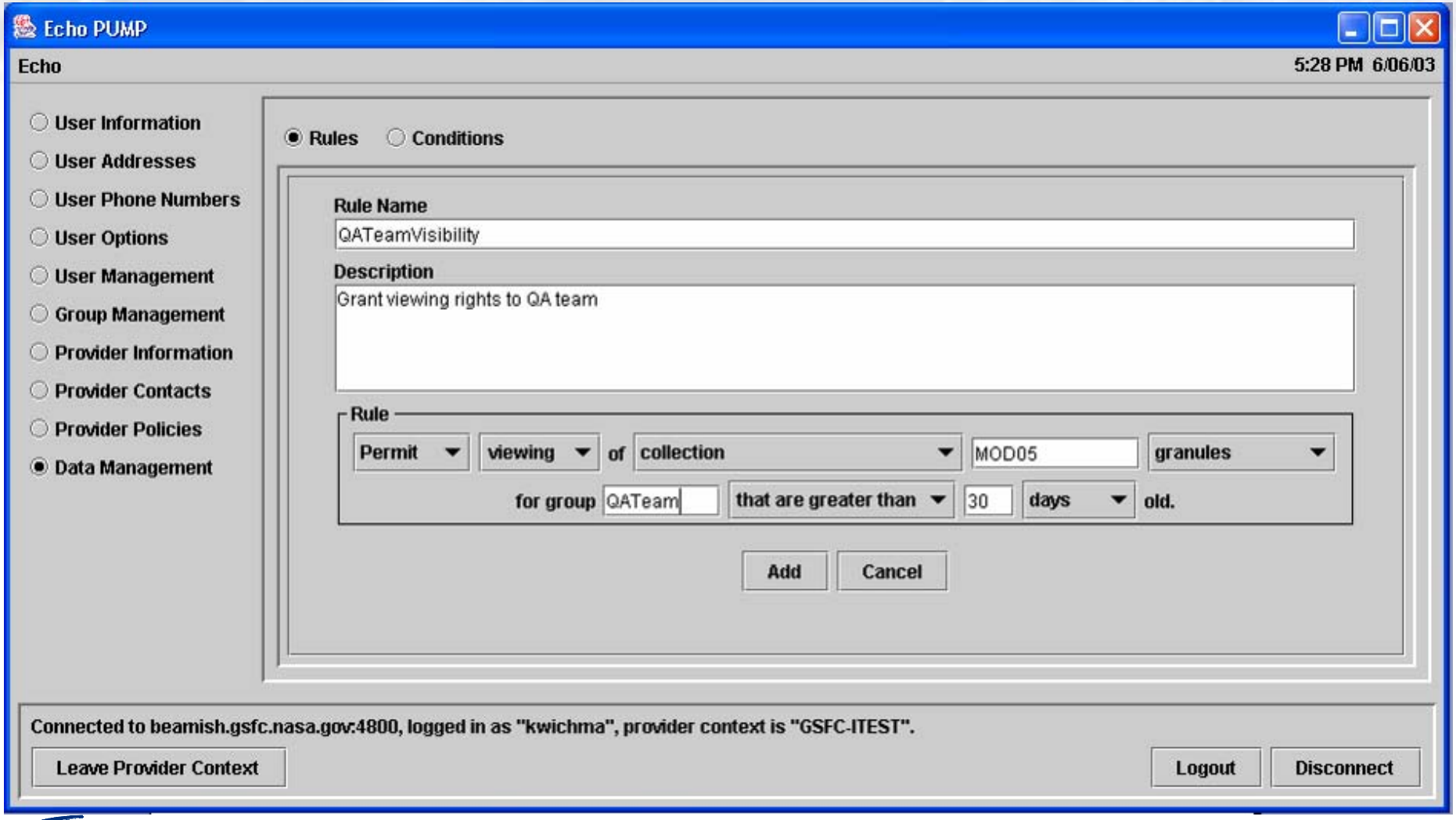
Data Management (ACL) – Provider Context



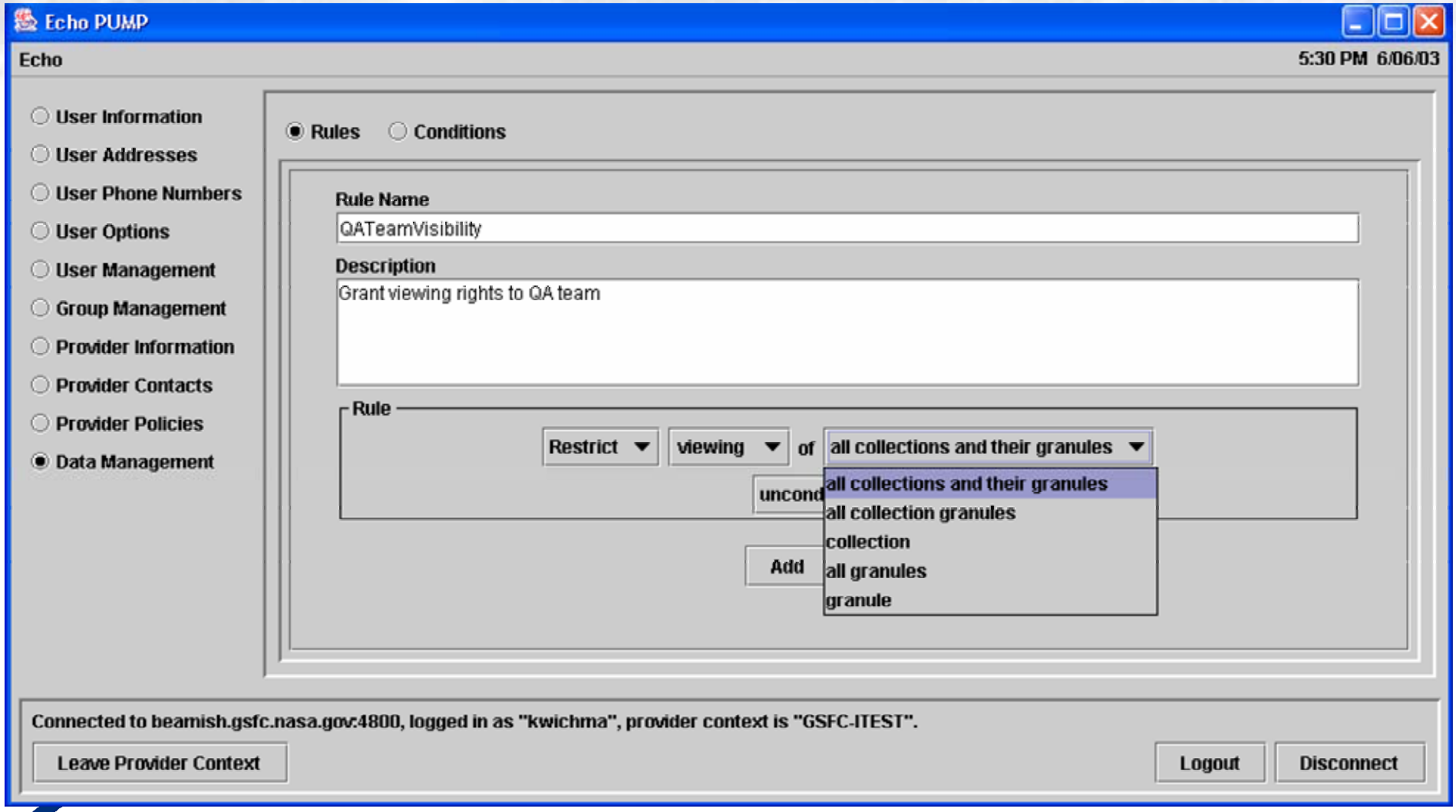
Create a Restriction – ACL – Provider Context



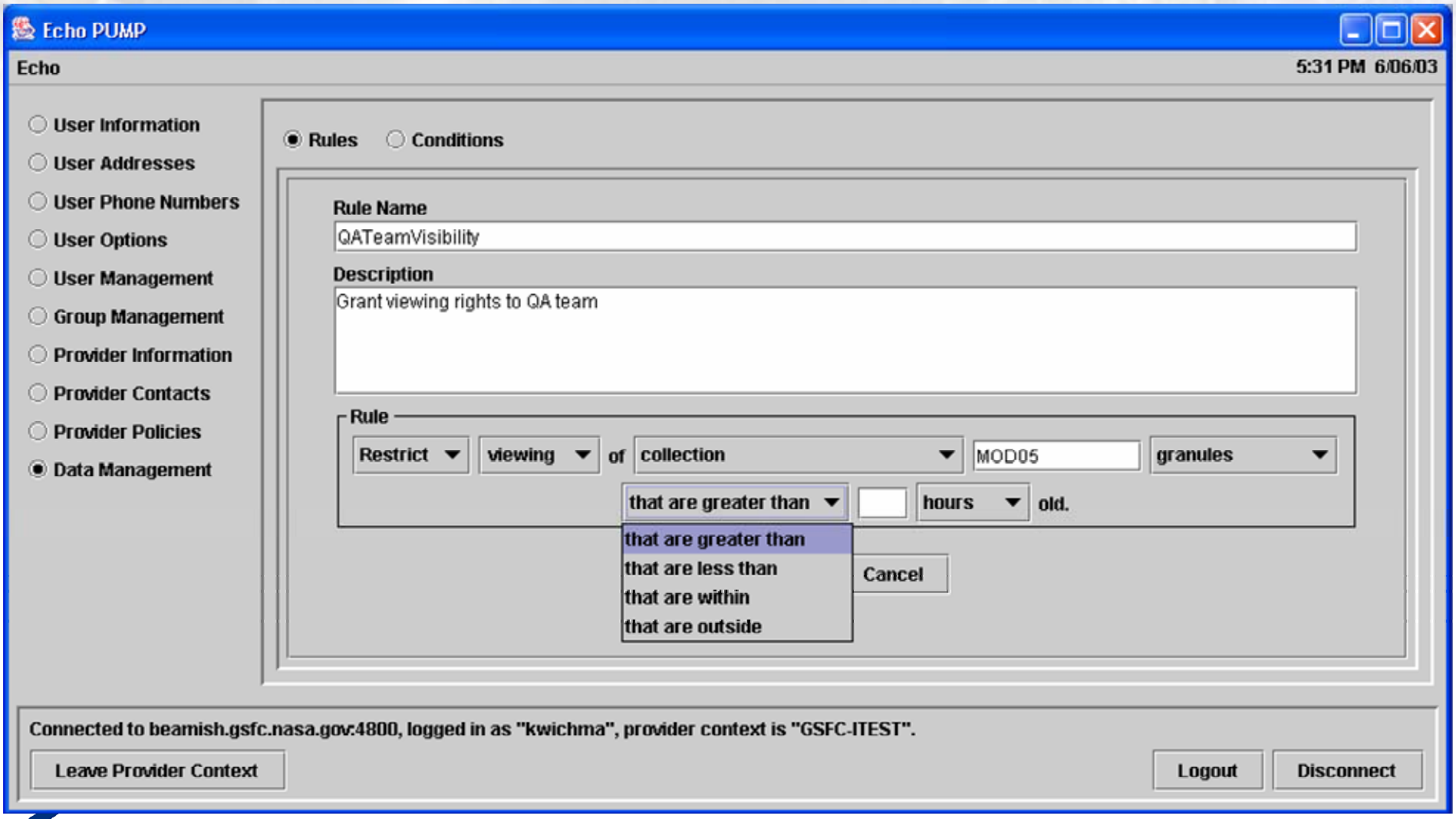
Create a Permission – ACL – Provider Context



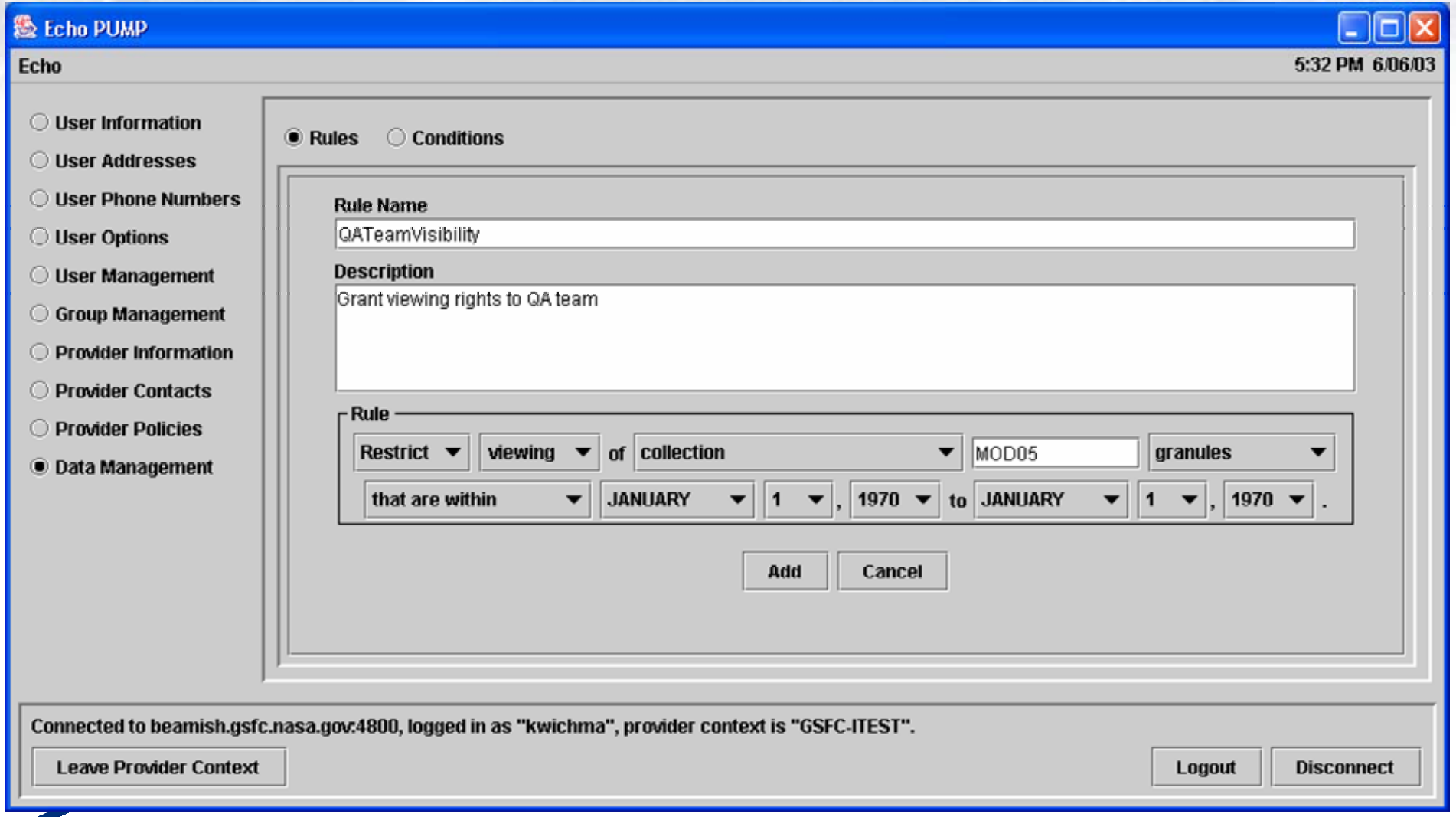
Restriction Options – ACL – Provider Context

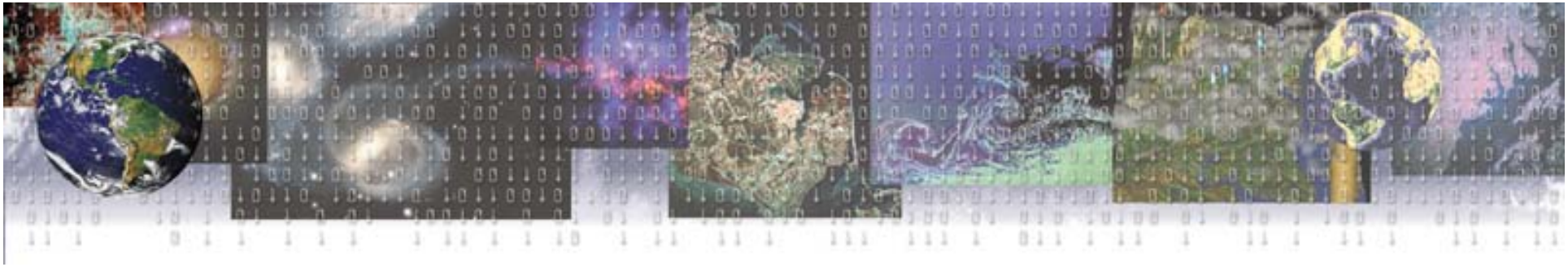


Condition Options – ACL – Provider Context



Fixed Temporal Range Restriction Example





Metadata Reconciliation

ECHO

NASA'S MISSION TO PLANET EARTH
EARTH PROBES
DATA INFORMATION SYSTEM
EOS
EARTH OBSERVING SYSTEM

Global Science & Technology, Inc.

NASA

Primary Issues

- ECHO represents data held at the data providers
- The update mechanism will be imperfect
- What are the mechanisms for determining when there is a difference between the data provider and ECHO?
- What are the mechanisms for correcting problems?

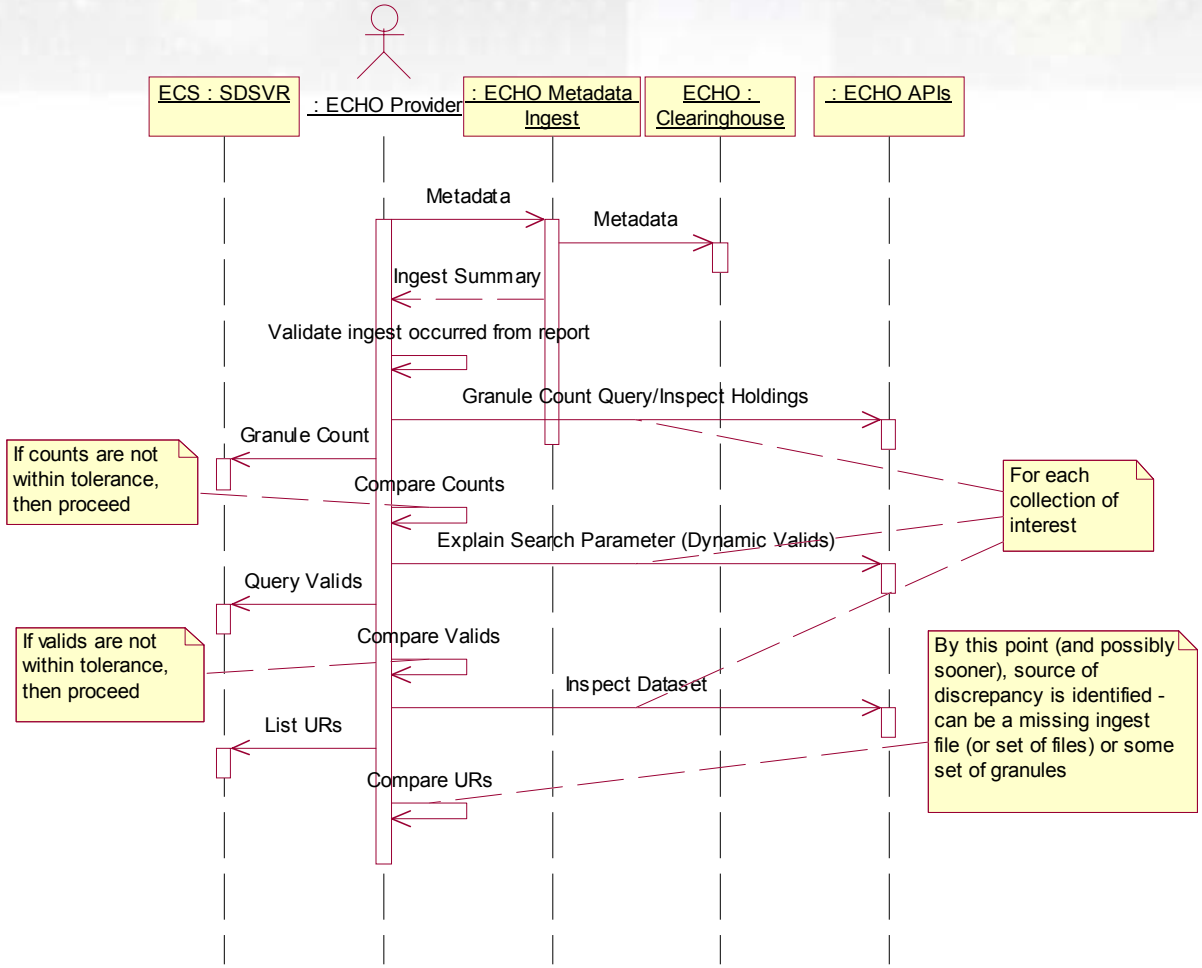
Discrepancy Detection

- **Monitor ingest reports from ECHO**
  - Were there errors?
  - Did ingest occur? (i.e. was there something sent and ECHO didn't get it?)
- **Compare counts**
  - Query ECHO for number of granules per collection and compare to data provider
    - Remember that there is always a delay in updating ECHO
- **Examine valids**
  - Query ECHO for extent of valids and compare to data provider
- **Compare holdings**
  - Query ECHO for a list of GranuleURs along with last update time and compare to data provider

Discrepancy Sources

- **ECHO is missing ingest files**
  - Export not run correctly
    - Addressed by proper configuration
  - Failure to deliver exported file
    - Addressed by monitoring delivery mechanism
  - Failure of ECHO to recognize and process exported file
    - Addressed by ECHO OPs monitoring ingest
- **ECHO is missing granules, or they are incorrect**
  - Problem with export program
    - Addressed by BMGT validation approaches
  - Problem with ECHO ingest
    - Addressed by internal and external tests of ECHO
  - Problem with ECHO query/results
    - Addressed by internal and external tests of ECHO

Discrepancy Detection



Approach for 4.5

- **Perform query yourself**

  -or-

- **Ask Operations Team to create a report for you**

Improvements for 5.0

- **ExplainSearchParameter transaction**
  - Sometimes referred to as dynamic valids
  - Used to determine what search attributes will return results given a query
  - i.e. What Platforms are valid if I search for an area equivalent to Maryland and limit my search to LPDAAC and GSFC?

5.5 Improvements

- **Inspection functions**
  - InspectHoldings
    - Lists all collections and number of granules in collection
    - Useful for high level check
    - Result returned inline
  - InspectDataset
    - Lists GranuleUR, Provider last update time, ECHO last update time
    - Uses simple delimited format
    - Can be returned inline, sent by email (compressed), or put on FTP site

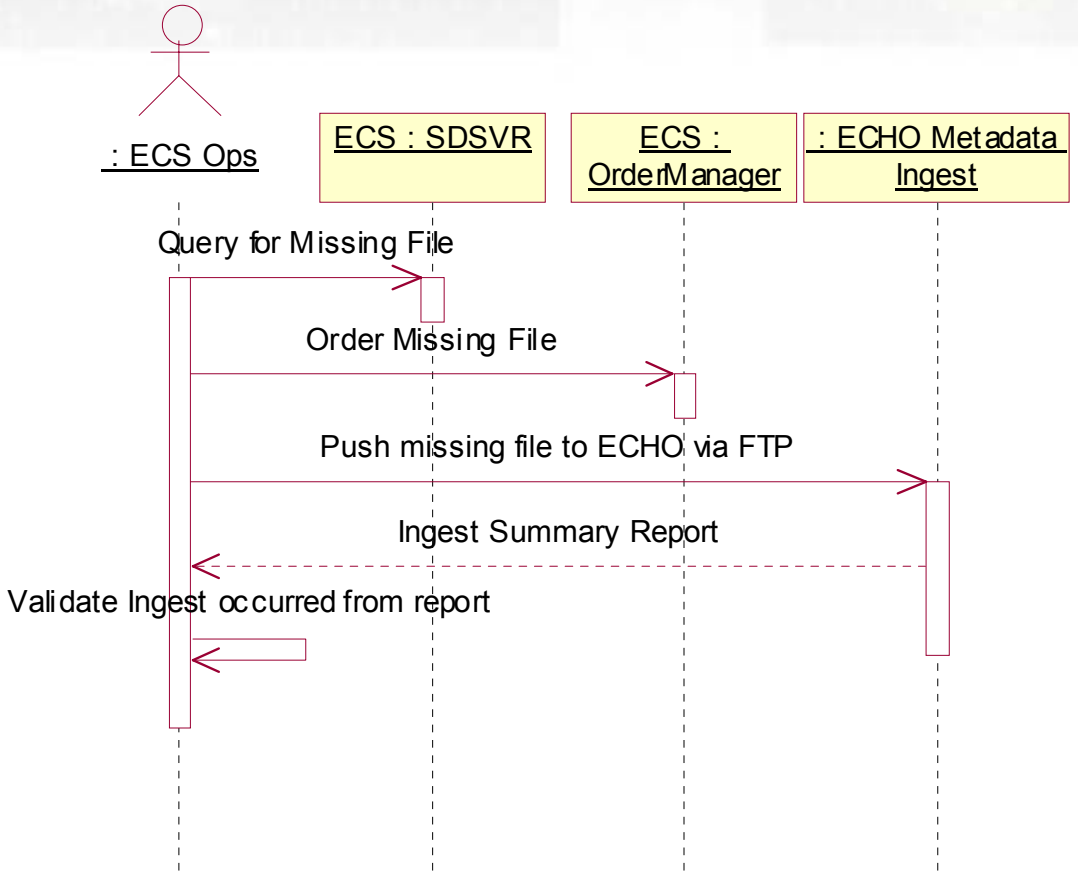
Reconciliation Options

- **Identify missing file and submit for ingest**
- **Create deletion list and submit**
- **Export missing data and submit for ingest**
- **Fix mapping issue and…**
  - Have operations fix in current archive, and adjust export/mapping/ingest appropriately

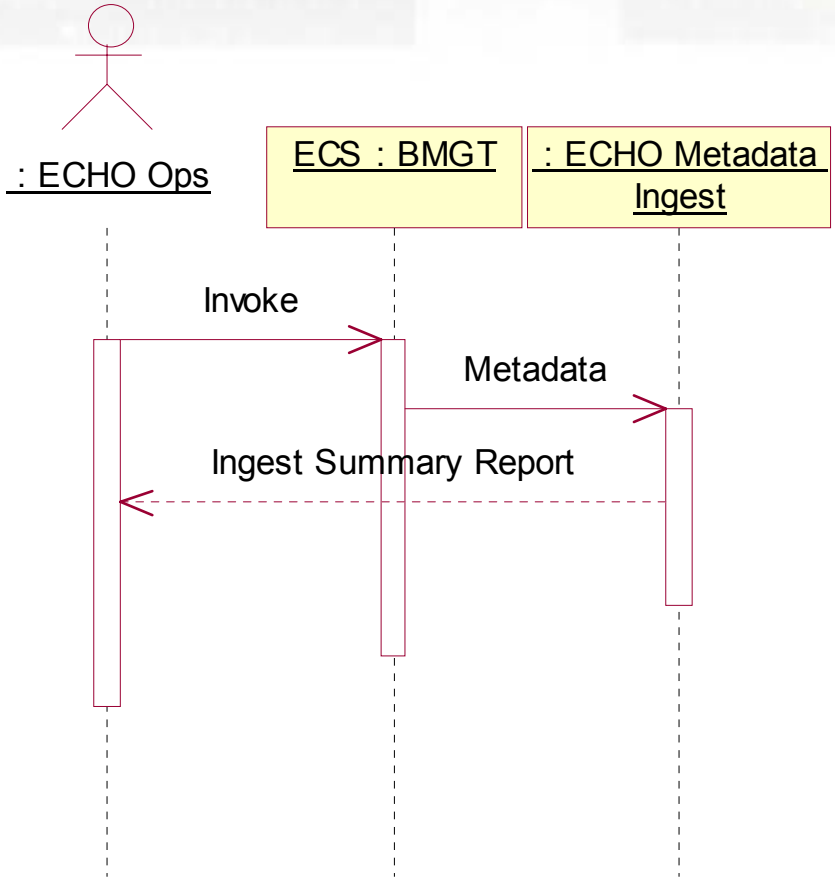
    -or-
  - Perform update of just changed metadata

    -or-
  - Delete all data and re-ingest

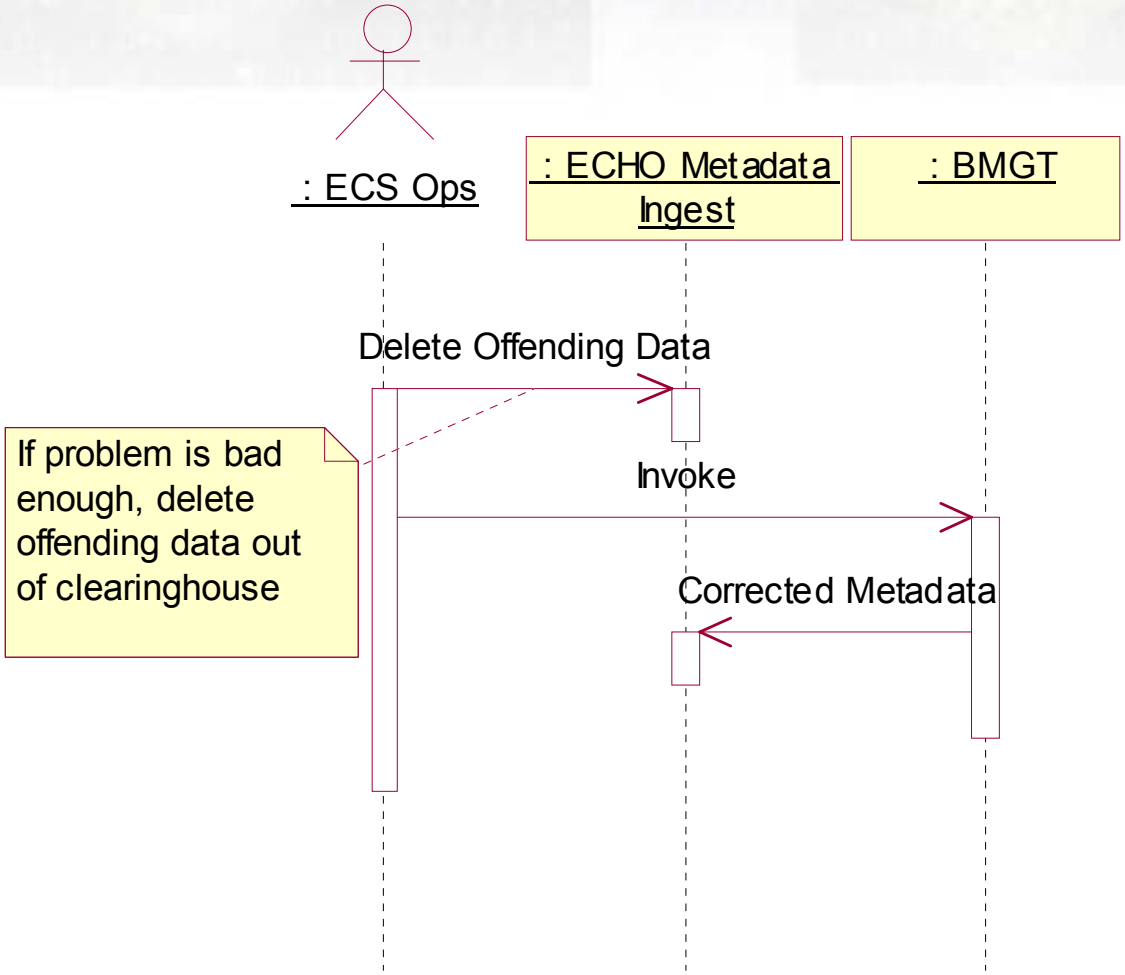
Metadata Reconciliation – Missing File(s)



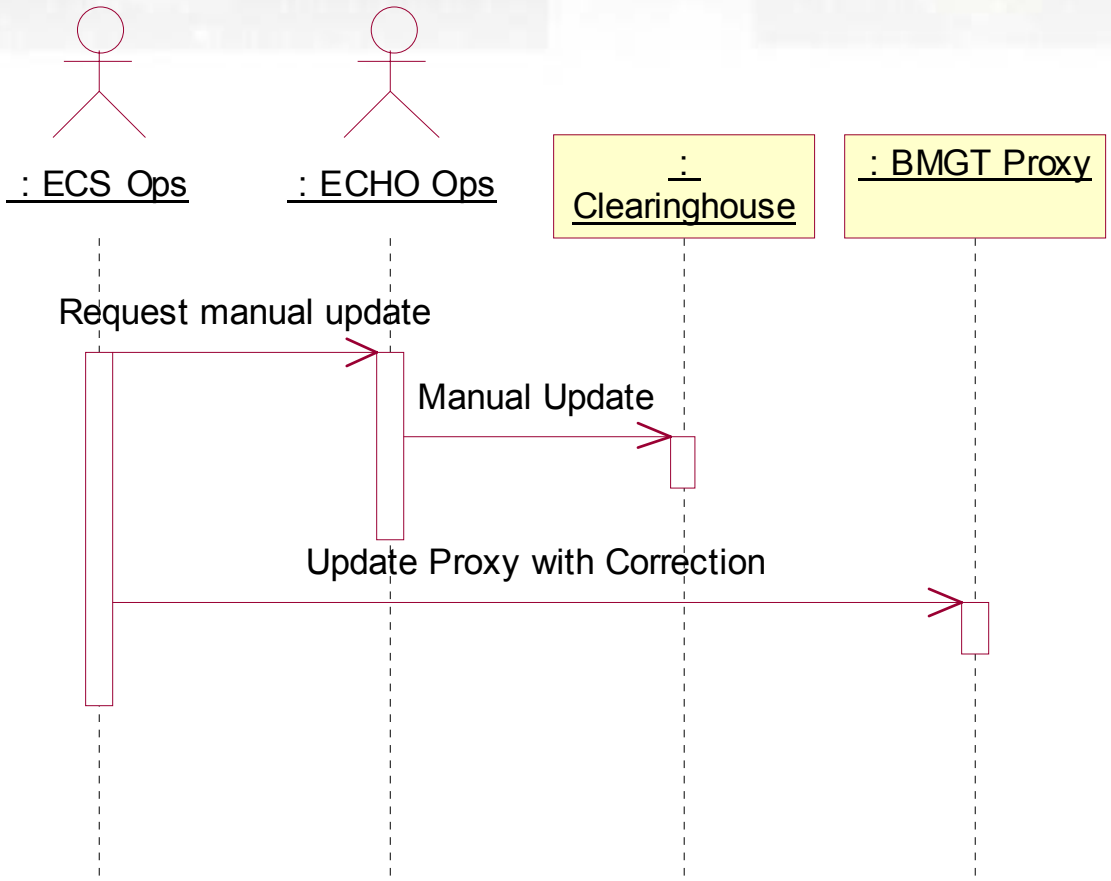
Metadata Reconciliation – Missing Granules

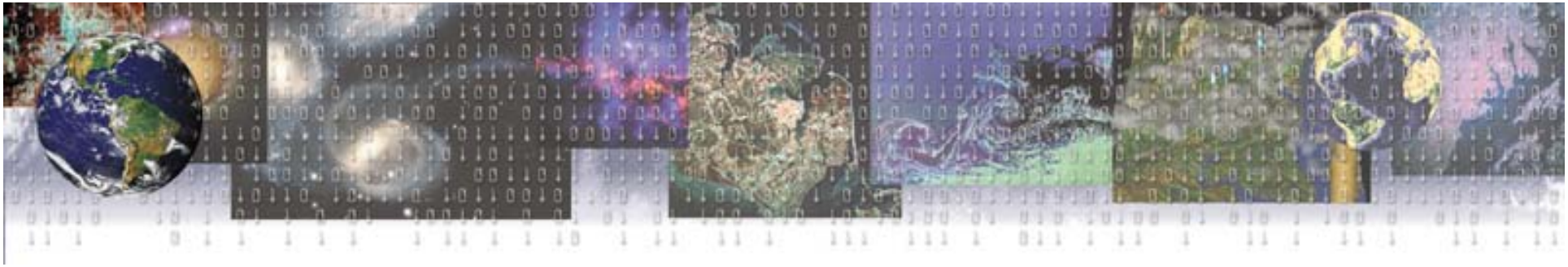


Metadata Reconciliation – Erroneous Granules

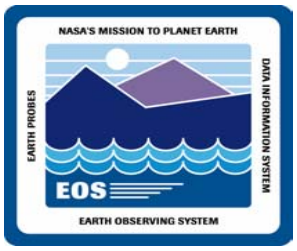


Metadata Reconciliation – Erroneous Granules (2)

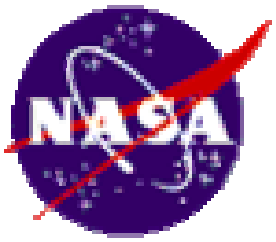




ECHO Test Systems

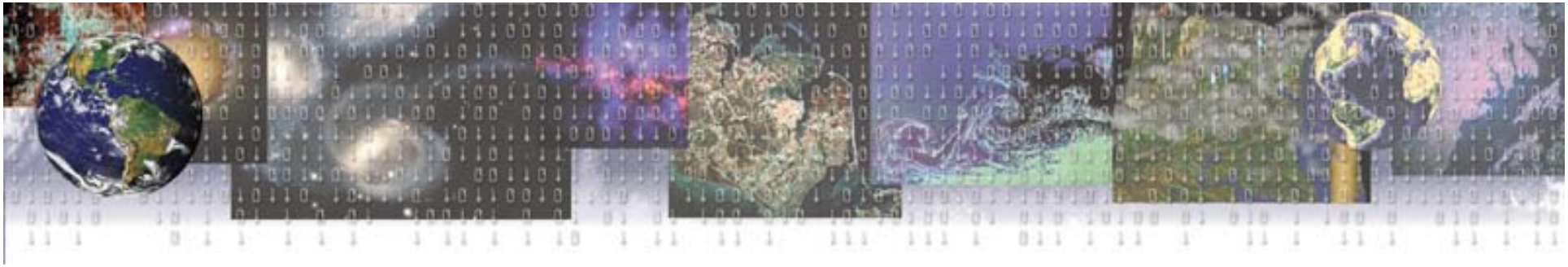


Global Science & Technology, Inc.



ECHO Instances

Previous Operational Version	Operational Version	External Test Version
Legacy System **Real Providers, Real Data, No ingest**	Operational System **Real Providers, Real Data, Ingest Activated**	External Test System **Dummy Providers, Test Data, Ingest Activated**
	Provider Test System **TSMode Providers, Real TSMode Data, Ingest Activated**	Client Test System **Dummy Providers, Test Data, No Ingest**
	Client Test System **Dummy Providers, Real Data + Test Data, No Ingest**	



RMA

ECHO

Global Science & Technology, Inc.

5 Sources of Downtime In ECHO

1. Hardware Error
2. Software Error
3. Reconfiguration
4. Backups
5. Logs fill up the disk space

Hardware Error

- **Occurs infrequently**
- **Has not happened since system went operational**
- **Requires hardware be fixed and/or replaced**
  - Typically a service call to Sun
  - Downtime of a day or two anticipated
- **Being addressed through introduction of clustered system**
  - This introduces redundancy in hardware with failover and load balancing capabilities.

Software Error

- **Occurs infrequently**
- **Has not happened since system went operational.**
  - Requires restarts of software.
- **Downtime is minutes once discovered (if someone is around to restart the system)**
- **Addressed through thorough testing and a robust environment**
  - Could be augmented by writing "nanny" script to watch system and restart it automatically.

Reconfiguration

- Occurs when certain system parameters have to be changed
- This is estimated to have happened on average about once a month.
- Downtime is minutes when it occurs.
- Being addressed by new versions of the system performing functions without need of system restart.

Backups

- **Occur weekly on Friday nights**
- **Downtime is roughly 4-5 hours once per week**
- **Being addressed by adding hot backup capability**
  - This should reduce downtime considerably
  - System will be limited to read-only functions for an estimated less than hour time period

Logs Fill Up Disk Space

- **Occurs when logs fill up allocated file system space**
- **This happened about 4-5 times since the system went operational and has now been corrected**
- **Downtime is minutes**
- **Addressed by reducing log file output, and changing how log files are managed to not require restart**
  - Some still require restart